

2011

Managing the Next Deluge: A Tax System Approach to Flood Insurance

Charlene Luke

Aviva Abramovsky

Follow this and additional works at: <https://opencommons.uconn.edu/cilj>

Recommended Citation

Luke, Charlene and Abramovsky, Aviva, "Managing the Next Deluge: A Tax System Approach to Flood Insurance" (2011). *Connecticut Insurance Law Journal*. 87.
<https://opencommons.uconn.edu/cilj/87>

CONNECTICUT INSURANCE LAW JOURNAL

Volume 18, Number 1
Fall 2011



University of Connecticut School of Law
Hartford, Connecticut

Connecticut Insurance Law Journal (ISSN 1081-9436) is published at least twice a year by the Connecticut Insurance Law Journal Association at the University of Connecticut School of Law. Periodicals postage paid at Hartford, Connecticut. Known office of publication: 55 Elizabeth Street, Hartford, Connecticut 06105-2209. **Printing location: Western Newspaper Publishing Company, 537 East Ohio Street, Indianapolis, Indiana 46204.**

Please visit our website at <http://www.insurancejournal.org> or see the final page of this issue for subscription and back issue ordering information.

Postmaster: Send address changes to Connecticut Insurance Law Journal, 55 Elizabeth Street, Hartford, Connecticut 06105-2209.

The Journal welcomes the submission of articles and book reviews. Both text and notes should be double or triple-spaced. Submissions in electronic form are encouraged, and should be in Microsoft™ Word™ version 97 format or higher. Citations should conform to the most recent edition of A UNIFORM SYSTEM OF CITATION, published by the Harvard Law Review Association.

It is the policy of the University of Connecticut to prohibit discrimination in education, employment, and in the provision of services on the basis of race, religion, sex, age, marital status, national origin, ancestry, sexual preference, status as a disabled veteran or veteran of the Vietnam Era, physical or mental disability, or record of such impairments, or mental retardation. University policy also prohibits discrimination in employment on the basis of a criminal record that is not related to the position being sought; and supports all state and federal civil rights statutes whether or not specifically cited within this statement.

Copyright © 2011 by the Connecticut Insurance Law Journal Association.

Cite as CONN. INS. L.J.

CONNECTICUT INSURANCE LAW JOURNAL

VOLUME 18

2011-2012

ISSUE 1

EDITORIAL BOARD 2011-2012

Editor-in-Chief

BETHANY L. DiMARZIO

Managing Editor

CAITLIN HOLT

Administrative Editor

RUBAIYAT MAHBOOB

Assistant Managing Editor

MICHAEL BARRESE

Lead Articles Editors

WYATT R. JANSEN
MARK LAVELLE

Notes & Comments Editors

RONALD X. JOHNSON
CAITLIN LOFTUS
TAYLOR MEEHAN
DONNY O'DAY

Executive Editors

CRAIG COOPER
KEEGAN MCCLURE
RYAN MCLAUGHLIN
GEOFF PERKINS
COLLEEN VELLTURO

Legal Abstracts Editor

MARTIN KRAEGEL

Research Editor

BRIAN R. SMITH

Competition Editor

ANDREA DiPAOLO

Symposium Editor

LESLEY CHUANG

Technology Editor

BRENDAN OWENS

JESSICA AURELIEN

TIMOTHY GOODWIN

Associate Editors

MATTHEW BERNARD

SUSAN MASTERS

DOUGLAS DALENA

WALTER MENJIVAR

MEMBERS

SCOTT BEWLEY
JEFFREY FLETCHER
DEVIN HARTLEY
KYLE LAMBRECHT
HAILEY PETRUCCELLI
THOMAS VAN ROTEN
ROBERT YAMARTINO

ROBERT CANNING
NICK GIULIANO
JEFFREY R. HOULE
JAY MCCORMICK
KRISTEN P. RAYMOND
ANDREW VEALE

AMANDA COLE
ANDREA GOMES
KELLY KIRBY
KERRY MOGENSEN
JAMES RINGOLD
BENJAMIN WEXLER

JOHN DOLAN
SHANE GOODRICH
TAMARA KRAMER
BRENDAN K. NELLIGAN
SUNAYANA SINGHANI
ALEXANDER WING

FACULTY ADVISOR

AVIVA ABRAMOVSKY

**UNIVERSITY OF CONNECTICUT
SCHOOL OF LAW**

**FACULTY AND OFFICERS OF ADMINISTRATION
FOR THE ACADEMIC YEAR 2011-2012**

OFFICERS OF ADMINISTRATION

Susan Herbst, Ph.D., *President, University of Connecticut*
Peter J. Nicholls, Ph.D., *Provost and Executive Vice President for Academic Affairs*
Jeremy Paul, J.D., *Dean, School of Law*
Darcy Kirk, J.D., *Associate Dean for Library and Technology*
Willajeanne F. McLean, J.D., *Associate Dean for Academic Affairs*
Ellen K. Rutt, J.D., *Associate Dean for Enrollment Management and Strategic Planning*
Susan Schmeiser, J.D., *Associate Dean for Student Life*
Steven Wilf, J.D., *Associate Dean for Research and Faculty Development*
Ann Crawford, M.Ed., *Assistant Dean for Finance and Administration*
Karen L. DeMeola, J.D., *Assistant Dean for Admissions and Student Finance*

FACULTY EMERITI

Phillip I. Blumberg, A.B., J.D., LL.D. (Hon.), *Dean and Professor of Law and Business, Emeritus*
John C. Brittain, B.A., J.D., *Professor of Law, Emeritus*
Clifford Davis, S.B., LL.B., *Professor of Law, Emeritus*
Nell J. Newton, B.A., J.D., *Dean and Professor of Law, Emerita*
Howard Sacks, A.B., LL.B., *Dean and Professor of Law, Emeritus*
George Schatzki, A.B., LL.B., LL.M., *Dean and Professor of Law, Emeritus*
Craig Shea, A.B., LL.B., LL.M., *Professor of Law, Emeritus*
Eileen Silverstein, A.D., J.D., *Professor of Law Emerita*
Lester B. Snyder, B.S., LL.B., LL.M., *Professor of Law, Emeritus*
Terry Tondro, A.B., LL.B., M.Phil., *Professor of Law, Emeritus*
Nicholas Wolfson, A.B., J.D., *George and Helen England Professor of Law, Emeritus*

FACULTY OF LAW

Jill Anderson, B.A., University of Washington; J.D., Columbia University; *William T. Golden Scholar Associate Professor of Law*
Paul Bader, B.A., Duke University; J.D., Mercer University Walter F. George School of Law; *Assistant Clinical Professor of Law*
Robin D. Barnes, B.A., J.D., State University of New York at Buffalo; LL.M., University of Wisconsin; *Professor of Law*
Jon Bauer, A.B., Cornell University; J.D., Yale University; *Clinical Professor of Law*
Loftus E. Becker, Jr., A.B., Harvard College; LL.B., University of Pennsylvania; *Professor of Law*
Bethany Berger, B.A., Wesleyan University; J.D., Yale University; *Professor of Law*
Robert Birmingham, A.B., J.D., Ph.D. (Econ.), Ph.D. (Phil.), University of Pittsburgh; LL.M., Harvard University; *Professor of Law*

Sara Bronin, B.A., University of Texas; M.Sc., University of Oxford (Magdalen College); J.D., Yale University; *Associate Professor of Law*

Deborah A. Calloway, B.A., Middlebury College; J.D., Georgetown University; *Professor of Law*

Marcia Canavan, B.A., University of California, Los Angeles; MPH., University of California, Los Angeles; J.D., Golden Gate University School of Law and University of Colorado School of Law; *Assistant Clinical Professor of Law*

Paul Chill, B.A., Wesleyan University; J.D., University of Connecticut; *Clinical Professor of Law*

John J. Chung, B.A., Washington University; J.D., Harvard University; *Visiting Professor of Law*

Stephen J. Cribari, B.A., St. Lawrence University; J.C.L., J.D., Catholic University; *Visiting Professor of Law*

Anne C. Dailey, B.A., Yale University; J.D., Harvard University; *Evangeline Starr Professor of Law*

Geoffrey Dellenbaugh, A.B., Princeton University; M.A., Stanford University; Ph.D., University of Pittsburgh; J.D., University of Chicago; *Associate Clinical Professor of Law*

Timothy H. Everett, B.A., M.A., Clark University; J.D., University of Connecticut; *Clinical Professor of Law*

Todd D. Fernow, B.A., Cornell University; J.D., University of Connecticut; *Professor of Law and Director, Clinical Programs*

Richard Michael Fischl, B.A., University of Illinois; J.D., Harvard University; *Professor of Law*

Eric J. Gouvin, B.A., Cornell University; M.P.A., Harvard University; J.D., LL.M., Boston University; *Visiting Professor of Law*

Hillary Greene, B.A., J.D., Yale University, *Professor of Law and Director, Intellectual Property Clinic*

Kaaryn Gustafson, A.B., Harvard University; Ph.D., J.D., University of California, Berkeley; *Professor of Law*

Mark W. Janis, A.B., Princeton University; B.A., M.A., Oxford University; J.D., Harvard University; *William F. Starr Professor of Law*

Dalie Jimenez, B.S., Massachusetts Institute of Technology; J.D. Harvard Law School; *Associate Professor of Law*

Richard S. Kay, A.B., Brandeis University; M.A.; Yale University; J.D., Harvard University; *Wallace Stevens Professor of Law*

Darcy Kirk, A.B., Vassar College; M.S., M.B.A., Simmons College; J.D., Boston College; *Professor of Law and Associate Dean for Library & Technology*

Peter R. Kochenburger, A.B., Yale University; J.D., Harvard University; *Associate Clinical Professor and Director of Graduate Programs*

Alan M. Kosloff, B.A., University of Connecticut; J.D., Georgetown University; *Director of Environmental Practice Clinic*

Lewis S. Kurlantzick, B.A., Wesleyan University; LL.B., Harvard University; *Zephaniah Swift Professor of Law*

James Kwak, A.B., Harvard College; Ph.D., University of California at Berkeley; J.D. Yale Law School; *Associate Professor of Law*

Alexandra Lahav, A.B., Brown University; J.D., Harvard University; *Professor of Law*

Lisa Laplante, A.B., Brown University; M.Ed., University of Massachusetts; J.D., New York University School of Law; *Visiting Assistant Professor and Richard D. Tulisano '69 Human Rights Fellow*

Leslie C. Levin, B.S.J., Northwestern University; J.D., Columbia University; *Professor of Law*

Diana L. Leyden, B.A., Union College; J.D., University of Connecticut; LL.M., Georgetown University; *Associate Clinical Professor of Law and Director, Tax Clinic*

Peter L. Lindseth, B.A., J.D., Cornell University; M.A., M. Phil, Ph.D., Columbia University; *Olimpiad S. Ioffe Professor of International and Comparative Law*

Joseph A. MacDougald, A.B., Brown University; M.B.A., New York University; J.D., University of Connecticut; M.E.M., Yale University; *Professor-in-Residence and Executive Director, Center for Energy and Environmental Law*

Hugh C. Macgill, B.A., Yale University; LL.B., University of Virginia; *Oliver Ellsworth Research Professor of Law and Dean Emeritus*

Neysun Mahboubi, A.B., Princeton University; J.D., Columbia University; *Research Fellow*

Jennifer Brown Mailly, A.B., Brown University; J.D., Ohio State University; *Assistant Clinical Professor of Law*

Miriam Marton, B.A., Oakland University; M.S.W., University of Michigan, J.D., Michigan State; *William R. Davis Clinical Teaching Fellow*

Ruth Mason, B.A., Columbia University; J.D., Harvard University; *Professor of Law*

Patricia A. McCoy, B.A., Oberlin College; J.D., University of California, Berkeley; *Professor of Law and Connecticut Mutual Chair in Insurance Law, Director of Insurance Law Center*

Barbara McGrath, B.A., Yale University; J.D., University of Connecticut; *Executive Director, Connecticut Urban Legal Initiative, Inc.*

Willajeanne F. McLean, B.A., Wellesley College; B.S., University of Massachusetts; J.D., Fordham University; LL.M., Free University of Brussels; *Professor of Law and Associate Dean for Academic Affairs*

George Mocsary, B.E., Cooper Union; M.B.A., University of Rochester; J.D., Fordham Law School; *Visiting Assistant Professor*

Thomas H. Morawetz, A.B., Harvard College; J.D., M.Phil., Ph.D., Yale University; *Tapping Reeve Professor of Law and Ethics*

Lily Neff, B.S.E.E., City College of New York; J.D., Pace University; *Assistant Clinical Professor of Law and Executive Director, Intellectual Property Clinic*

R. Kent Newmyer, B.A., Doane College; Ph.D., University of Nebraska; *Professor of Law and History*

Angel R. Oquendo, A.B., M.A., Ph.D., Harvard University; J.D., Yale University; *George J. and Helen M. England Professor of Law*

Leonard Orland, B.A., Rutgers University; LL.B., University of Pennsylvania; *Oliver Ellsworth Research Professor of Law*

Sachin Pandya, B.A., University of California, Berkeley; M.A., Columbia University; J.D., Yale University; *Associate Professor of Law*

Richard W. Parker, A.B., Princeton University; J.D., Yale University; D.Phil., Oxford University; *Professor of Law*

Jeremy Paul, A.B., Princeton University; J.D., Harvard University; *Dean and Thomas F. Gallivan, Jr. Professor of Real Property Law*

Hon. Ellen Ash Peters, B.A., Swarthmore College; LL.B., Yale University; LL.D., Yale University; University of Connecticut; et al.; *Visiting Professor of Law*

Richard D. Pomp, B.S., University of Michigan; J.D., Harvard University; *Alva P. Loiselle Professor of Law*

Jessica S. Rubin, B.S., J.D., Cornell University; *Assistant Clinical Professor of Law*

Susan R. Schmeiser, A.B., Princeton University; J.D., Yale University; Ph.D., Brown University; *Professor of Law and Associate Dean for Student Life*

Peter Siegelman, B.A., Swarthmore College; M.S.L., Ph.D., Yale University; *Roger Sherman Professor of Law*
 Gordon Silverstein, B.A., Cornell University; Ph.D., Harvard University; *Visiting Professor*
 James H. Stark, A.B., Cornell University; J.D., Columbia University; *Professor of Law and Director, Mediation Clinic*
 Martha Stone, B.A., Wheaton College; J.D., LL.M., Georgetown University; *Director, Center for Children's Advocacy*
 Kurt A. Strasser, B.A., J.D., Vanderbilt University; LL.M., J.S.D., Columbia University; *Phillip I. Blumberg Professor of Law*
 Colin C. Tait, A.B., Cornell University; LL.B., Yale University; *Oliver Ellsworth Research Professor of Law*
 Stephen G. Utz, B.A., Louisiana State University; J.D., University of Texas; Ph.D., Cambridge University; *Professor of Law*
 Carol Ann Weisbrod, J.D., Columbia University; *Ellen Ash Peters Professor of Law*
 Robert Whitman, B.B.A., City College of New York; J.D., Columbia University; LL.M., New York University; *Professor of Law*
 Steven Wilf, B.S., Arizona State University; Ph.D., J.D., Yale University; *Joel Barlow Professor of Law and Associate Dean for Research and Faculty Development*
 Richard A. Wilson, BSc., Ph.D., London School of Economics and Political Science; *Professor of Law*

ADJUNCT FACULTY OF LAW

Peter G. Austin, B.S., University of Connecticut; M.S., University of Hartford; *Adjunct Professor of Law*
 Morris W. Banks, A.B., Dartmouth College; LL.B., Columbia University; LL.M., New York University; *Adjunct Professor of Law*
 Anne Davis Barry, B.S., University of Connecticut; M.S., Union College; J.D., University of Connecticut; *Lecturer in Law*
 James W. Bergenn, B.A., Catholic University; J.D., Columbia University; *Adjunct Professor of Law*
 Hon. David M. Borden, B.A., Amherst College; LL.B., Harvard University; *Adjunct Professor of Law*
 Leonard Boyle, B.A., University of Hartford; J.D., University of Connecticut; *Lecturer in Law*
 John Buchanan, A.B., Princeton, B.A., Oxford; J.D., Harvard; *Lecturer in Law*
 Michael A. Cantor, B.S., J.D., University of Connecticut; *Adjunct Professor of Law*
 Robert F. Carter, A.B., Harvard; M.A., J.D., Yale University; *Lecturer in Law*
 H. Wayne Carver II, Sc.B., Brown University; M.D., Brown University; *Senior Lecturer in Law and Forensic Science*
 Robert Casale, B.A., Southern Connecticut State University; M.A., Harvard University; J.D., John Marshall Law School; *Lecturer in Law*
 Audrey R. Chapman, B.A., Wellesley College; M.A., Ph.D., Columbia University; M.Div., New York Theological Seminary; S.T.M., Union Theological Seminary; *Lecturer in Law*
 James N. Cohen, B.A., J.D., Harvard; *Lecturer in Law*
 Gary Collins, B.A., State University of New York, Buffalo; J.D., Vanderbilt University; *Lecturer in Law*
 Dean M. Cordiano, B.A., State University of New York at Binghamton; J.D., Duke University; *Lecturer in Law*

Michelle Cruz, B.A., M.A., Mount Holyoke College; J.D., University of Connecticut; *Lecturer in Law*

Richard Cummings, B.A., University of Minnesota; LL.B., University of Minnesota Law School; M.B.A. University of Minnesota School of Business; *Visiting Assistant Professor*

Renee M. Dailey, B.A., J.D., *University of Connecticut; Lecturer in Law*

Nora Dannehy, B.A., Wellesley College; J.D., Harvard University; *Lecturer in Law*

John G. Day, A.B., Oberlin College; J.D., Case Western Reserve University; *Professor in Residence*

Robert DeCrescenzo, B.A., University of Connecticut; M.P.A., University of Connecticut; J.D., University of Connecticut; *Lecturer in Law*

Raymond DeMeo, B.A., Williams College; J.D., University of Connecticut; *Lecturer in Law*

Karen DeMeola, B.A., University of Connecticut; J.D., University of Connecticut; *Lecturer in Law*

Hon. Alexandra DiPentima, A.B., Princeton University; J.D., University of Connecticut; *Lecturer in Law*

Mark Dubois, B.A., College of the Holy Cross; J.D., University of Connecticut; *Lecturer in Law*

Gerald P. Dwyer, A.B., Colgate University; J.D., Gonzaga University; LL.M., Georgetown University; *Lecturer in Law*

Michael Eisele, B.S., U.S. Merchant Marine Academy; J.D., Georgetown University; *Lecturer in Law*

Thomas Farrish, B.A., J.D., University of Connecticut; *Lecturer in Law*

Rebecca Flanagan, B.A., M.A., University of Connecticut; J.D., University of North Carolina; *Lecturer in Law*

Evan D. Flaschen, B.A., Wesleyan University; J.D., University of Connecticut; *Adjunct Professor of Law*

Charles W. Fortune, B.A., University of Rochester; J.D., University of Connecticut; *Lecturer in Law*

Michael A. Gailor, B.A., Cornell University; J.D., University of Connecticut; *Lecturer in Law*

William Goddard, B.A., M.B.A., Dartmouth; J.D., University of Connecticut; *Lecturer in Law*

Ira H. Goldman, B.A., Cornell University; J.D., Yale University; *Adjunct Professor of Law*

Daniel L. Gottfried, B.A., J.D., University of Connecticut; *Lecturer in Law*

James Greenfield, A.B. Brown University; M.B.A., Columbia University; J.D., Columbia University School of Law; *Lecturer in Law*

Andrew S. Groher, B.A., University of Virginia; J.D., University of Connecticut; *Adjunct Professor of Law*

Christopher (Kip) Hall, B.A., Dartmouth College; B.A., Oxford University; J.D., University of Chicago Law School; *Lecturer in Law*

Albert B. Harper, B.A., University of Texas; J.D., Ph.D., University of Connecticut; *Adjunct Professor of Law*

John Harris, B.S.E., Princeton; J.D., University of Wisconsin; *Adjunct Professor of Law*

Jonathan A. Harris, B.S., Duke University; M.B.A., University of Connecticut School of Business; J.D., University of Connecticut; *Lecturer in Law*

Elizabeth Holland, B.A., Providence College; M.A., University of London; J.D. Suffolk University Law School; *Lecturer in Law*

Wesley Horton, B.A., Harvard University; J.D., University of Connecticut; *Adjunct Professor of Law*

John Houlihan, B.A., Providence College; J.D., St. John's University; *Lecturer in Law*

Hugh F. Keefe, B.S., M.S., Quinnipiac College; J.D., University of Connecticut; *Lecturer in Law*

Jocelyn Kennedy, B.A., University of New Hampshire; M.L.I.S., University of Washington; J.D., University of New Hampshire School of Law; *Lecturer in Law*

Nancy Kennedy, B.A., University of Massachusetts; J.D., University of Connecticut; *Lecturer in Law*

Daniel Klau, B.A., University of California, San Diego; J.D., Boston University; *Lecturer in Law*

Charles H. Klippel, B.A., M.Ed., M.P.H., J.D., Harvard University; *Lecturer in Law*

Amy Kokoski, B.S., Rensselaer Polytechnic Institute; J.D., University of Connecticut; *Lecturer in Law*

Barry Kramer, B.Ch.E., Rensselaer Polytechnic Institute; J.D., New York University; *Lecturer in Law*

Carll Ladd, B.S., Bates College; Ph.D., University of Connecticut; *Senior Lecturer in Law and Forensic Science*

John H. Lawrence, Jr., B.S., Washington and Lee University; J.D., University of Virginia; *Lecturer in Law*

Zita Lazzarini, B.A. University of California at Berkeley; J.D., University of California Hastings College of Law; M.P.H., Harvard School of Public Health; *Lecturer in Law*

Henry C. Lee, B.S., John Jay College of Criminal Justice; M.S., Ph.D., New York University; Dr.Sci. (Hon.), University of New Haven; Dr.Hum. (Hon.), St. Joseph College; *Adjunct Professor of Law*

Erik T. Lohr, B.S., Thomas A. Edison State College; J.D., University of Connecticut; *Lecturer in Law*

James Lotstein, B.S., Northwestern University; J.D., University of Connecticut, *Lecturer in Law*

Thomas S. Marrion, A.B., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*

John Mastropietro, B.A., University of Connecticut; J.D., Western New England College; *Lecturer in Law*

Elsie Mata, B.A., St. John's College; J.D., University of Michigan Law School; *Lecturer in Law*

Virginia M. Maxwell, B.Sc., Liverpool University; D.Phil., The Queen's College; *Senior Lecturer in Law and Forensic Science*

Paul McCary, B.A. University of Connecticut; J.D., Western New England College School of Law; *Lecturer in Law*

Dwight Merriam, B.A., University of Massachusetts, M.R.P., University of North Carolina, J.D., Yale University; *Lecturer in Law*

Joseph Mirrione, B.A., Marist College; J.D., Vermont Law School; *Lecturer in Law*

Derek L. Mogck, B.A., Gordon College; M.S., University of Massachusetts; J.D., University of Connecticut; *Lecturer in Law*

Umar F. Moghul, B.A., M.A., University of Pennsylvania; J.D., Temple University; *Lecturer in Law*

Thomas B. Mooney, B.A., Yale University; J.D., Harvard University; *Adjunct Professor of Law*

Ronald T. Murphy, B.S., Springfield College; J.D., University of Connecticut; *Lecturer in Law*

Henry F. Murray, B.A., Yale University; J.D. University of Connecticut; *Lecturer in Law*

Mary R. Norris, B.S., University of Cincinnati; J.D., Georgetown University; *Lecturer in Law*

Andrew J. O'Keefe, B.S., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*

Cornelius O'Leary, B.A., Williams College; M.A., Trinity College; J.D., University of Connecticut; *Adjunct Professor of Law and Mark A. Weinstein Clinical Teaching Fellow*

Jonathan B. Orleans, B.A., Hampshire College; M.P.A., Princeton University; J.D., New York University; *Lecturer in Law*

Rosemarie Paine, B.S., Southern Connecticut State University; J.D., University of Connecticut; *Lecturer in Law*

Humbert J. Polito, Jr., A.B., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*

Elliott B. Pollock, A.B., Columbia College; LL.B., Columbia Law School; *Lecturer in Law*

Hon. Elliot D. Prescott, B.A., University of Massachusetts; J.D., University of Connecticut; *Lecturer in Law*

Renee C. Redman, B.A., Michigan State University; J.D., Brooklyn Law School; *Lecturer in Law*

Leah Reimer, B.S., J.D., Ph.D., University of Connecticut; *Lecturer in Law*

Roger Reynolds, B.A., Macalester College; J.D., New York University; *Lecturer in Law*

Barbara Reznar, B.A., Gettysburg College; J.D., Washington and Lee University; *Lecturer in Law*

Hon. Chase Rogers, B.A., Stanford University; J.D., Boston University School of Law; *Lecturer in Law*

William M. Rubenstein, B.A., Montclair State University; J.D. Western New England College; *Lecturer in Law*

Morgan Paul Rueckert, B.A., J.D., University of Connecticut; *Lecturer in Law and Mark A. Weinstein Clinical Teaching Fellow*

Ellen Rutt, B.A., University of Notre Dame; M.A., Fairfield University; M.B.A., American Graduate School of International Management; J.D., University of Connecticut; *Lecturer in Law*

Patrick J. Salve, B.S., J.D., University of Pennsylvania; *Adjunct Professor of Law*

Hon. Michael R. Sheldon, A.B., Princeton University; J.D., Yale University; *Adjunct Professor of Law*

Leah Lopez Schmalz, B.A./B.ID., University of Louisiana, Lafayette; J.D., Pace University School of Law; *Lecturer in Law*

Sandra L. Sherlock-White, B.A., Central Connecticut State College; J.D., Western New England College; *Lecturer in Law*

Jay E. Sicklick, B.A., Colgate University; J.D., Boston College; *Adjunct Professor of Law*

Richard L. Sigal, B.A., Yale University; J.D., University of Chicago; *Lecturer in Law*

Douglas Simpson, A.B., Dartmouth College; J.D., University of Connecticut; *Lecturer in Law*

Jason A. Smith, B.A., University of Tulsa; M.T.S., Harvard Divinity School; J.D., Northeastern University School of Law; *Lecturer in Law*

Ruth Sovronsky, B.A., Queens College, City University of New York; J.D., St. John's University School of Law; *Lecturer in Law*

Royal J. Stark, B.A., J.D., University of Connecticut; *Lecturer in Law*

Joshua Stein, B.A., Yale University; J.D. University of Connecticut; *Lecturer in Law*

Hon. Mark H. Taylor, B.A., Drew University; J.D., University of Connecticut; *Lecturer in Law*

James R. Turcotte, B.A., Boston University; J.D., Western New England College; *Lecturer in Law*

Christopher F. Wanat, B.A., Connecticut College; J.D., Dickinson School of Law; *Lecturer of Law*
Rosalind Wiggins, B.A., Yale College; J.D., Stanford Law School; *Lecturer in Law*
William Wilcox, B.A., University of Connecticut; J.D., Suffolk University; *Lecturer in Law*
Kenneth Burton Zercie, A.S., University of New Haven; B.S., University of New Haven; M.S., University of New Haven; *Senior Lecturer in Law and Forensic Science*
Jeremy (Remy) Zimmermann, A.B., Brown University; LL.B., Georgetown University; *Lecturer in Law*

CONNECTICUT INSURANCE LAW JOURNAL

VOLUME 18 2011-2012 NUMBER 1

CONTENTS

ARTICLES

MANAGING THE NEXT DELUGE:
A TAX SYSTEM APPROACH TO
FLOOD INSURANCE *Charlene Luke & Aviva Abramovsky* 1

A CONCURRENT MESS AND A CALL FOR
CLARITY IN FIRST-PARTY PROPERTY
INSURANCE COVERAGE ANALYSIS *Mark M. Bell* 73

INSURANCE RATES REGULATION
IN COMPARISON WITH OPEN
COMPETITION *Angelo Borselli* 109

ERISA: REMEDIES, PREEMPTION AND
THE NEED FOR MORE STATE
REGULATORY OVERSIGHT *Jillian Redding, Esq.* 169

SYMPOSIUM

KEYNOTE ADDRESS: ACTUARIAL
LITIGATION: HOW BIG STATISTICS
CAN HELP SOLVE BIG CASES *Kenneth Feinberg* 221

A NORMATIVE EVALUATION OF
ACTUARIAL LITIGATION *Robert G. Bone* 227

REEVALUATING COMPLEX MEDIATION GENERALIZATIONS	<i>Edward Brunet</i>	279
---	----------------------	-----

PROBABILITY SAMPLING IN LITIGATION	<i>Joseph B. Kadane</i>	297
---------------------------------------	-------------------------	-----

NOTES AND COMMENTARIES

IMPERMISSIBLE WINDFALLS?: UNEMPLOYMENT INSURANCE, BACK PAY, AND THE TWO CLASSES OF TITLE VII PLAINTIFFS	<i>Wyatt R. Jansen</i>	307
--	------------------------	-----

LESSONS FROM THE PRICE-ANDERSON NUCLEAR INDUSTRY INDEMNITY ACT FOR FUTURE CLEAN ENERGY COMPENSATORY MODELS	<i>Taylor Meehan</i>	339
---	----------------------	-----

MANAGING THE NEXT DELUGE: A TAX SYSTEM APPROACH TO FLOOD INSURANCE

CHARLENE LUKE[†] & AVIVA ABRAMOVSKY[‡]

The National Flood Insurance Program (NFIP) has fallen short in fulfilling its promise as a social safety net for flood loss victims. In place of the NFIP, this Article proposes a mandatory social insurance plan that would harness the strengths of the federal taxing authority to provide basic relief for flood losses occurring at an individual's primary residence. Any plan for addressing flood loss must navigate hotly debated, competing views about government intervention, redistribution, private markets, environmental protection, and property rights. This Article argues that government intervention in flood loss relief is inevitable, at least in the foreseeable future, and that the focus of that intervention should be on the ex ante provision of a social safety net. The program proposed in this Article is also intended to provide additional levers for addressing the complexities of flood loss, including the reduction of negative environmental externalities, and to provide the impetus needed for harmonizing existing tax provisions and grant programs.

I. INTRODUCTION

Early on the morning of August 30, 2015, the life of Alice and her son will change forever when floodwater rips through the ground floor apartment rented by Alice. Miraculously, Alice will have sufficient

[†] Associate Professor, University of Florida Levin College of Law.

[‡] Associate Professor, Syracuse University College of Law.

The authors thank Yariv Brauner, Jonathan R. Cohen, Patricia E. Dilley, Mark Fenster, Alyson C. Flournoy, Michael K. Friel, Carlo Garbarino, Thomas R. Hurst, Robert H. Jerry, II, David Cay Johnston, Shani M. King, Christine A. Klein, Leandra Lederman, Tom Lin, Martin McMahon, Jr., D. Daniel Sokol, and Lee-ford Tritt for their suggestions on the project. An early version of the project was presented at the 2009 Law & Society Annual Meeting, and we thank the participants at that event. We also thank research assistants Justin Bryan, Christy Corrao, Rachel Espey, James Lang, Amber Morris, Gustavo Morles & Sheila Welch. Any errors are our own.

warning of the imminent collapse of a dam that she and her son will be able to escape with their lives.¹ Many of the personal possessions that will be destroyed in the disaster are irreplaceable — the first baby tooth lost by her son and saved by Alice, the family photographs that Alice never has had the time or money to digitize and upload to the cloud, the souvenirs Alice purchased on a road trip taken many years ago when times were better. Alice will, however, be able to take some comfort in the knowledge that with each paycheck she has received over the past three years, she has been participating in a national flood loss security plan — a plan that will now help her in making a dignified fresh start.

If, however, the National Flood Insurance Program (NFIP) continues on its present course, the outcome for Alice may well be very different. Without new legislation, the program will not even exist in 2015; in 2010 the program briefly lapsed,² and in 2011 the program has been extended for multiple short-term periods with the most recent extension ending on December 23, 2011.³ Even if Congress acts to extend the current version of the NFIP, Alice will almost certainly not have purchased flood insurance because of the low participation rates associated with the NFIP. Instead, Alice will likely be scrambling for *ad hoc*, piecemeal post-disaster assistance.⁴ She may think back to the news coverage of ten years before⁵

¹ See Henry Fountain, *Danger Pent Up Behind Aging Dams*, N.Y. TIMES, Feb. 22, 2011, at D1, available at <http://www.nytimes.com/2011/02/22/science/22dam.html?pagewanted=all> (“[O]f the nation's 85,000 dams, more than 4,400 are considered susceptible to failure . . .”).

² See Rebecca Mowbray, *Lapses in National Flood Insurance Program Bring Policy Renewals to a Halt*, NOLA.COM (June 30, 2010, 6:53PM), http://www.nola.com/business/index.ssf/2010/06/lapses_in_national_flood_insur.html.

³ As of December 21, 2011. See Resolution Making Further Continuing Appropriations, Pub. L. No. 112-68 (extension through Dec. 23, 2011); Resolution Making Further Continuing Appropriations, Pub. L. No. 112-67, 125 Stat. 769 (extension through Dec. 17, 2011); Consolidated & Further Continuing Appropriations Act, 2012, Pub. L. No. 112-55 § 101, 125 Stat. 552 (extension through Dec. 16, 2011); Continuing Appropriations Act, 2012, Pub. L. No. 112-36 § 130, 125 Stat. 386 (extension through Nov. 18, 2011); Continuing Appropriations Act, 2012, Pub. L. No. 112-33, 125 Stat. 363 (extension through Oct. 4, 2011); National Flood Insurance Program Reextension Act of 2010, Pub. L. No. 111-250, 124 Stat. 2630 (to be codified as amended at 42 U.S.C. § 4026) (extension through Sept. 30, 2011). Additional extensions have been proposed, including one that would extend the NFIP through May 2012. See H.R. 3628, 112th Cong. (2011); S. 1548, 112th Cong. (2011).

⁴ See Christine A. Klein & Sandra B. Zellmer, *Mississippi River Stories: Lessons from a Century of Unnatural Disasters*, 60 SMU L. REV. 1471, 1473

and realize that she has become trapped in her own version of Hurricane Katrina.

Flood losses are only likely to escalate in the coming years.⁶ Before the next massive flood occurs⁷ — indeed before the next flood that devastates an individual life occurs — Congress should enact a new program for flood loss relief that provides a better social safety net than the current NFIP. This Article suggests a mandatory social insurance plan that

(2007) (“Too often, those who suffer most are the poorest members of society. . . .”). Cf. Saul Levmore & Kyle D. Logue, *Insurance Against Terrorism—And Crime*, 102 MICH. L. REV. 268, 277 (2003) (predicting that “public and charitable relief will more likely be forthcoming if there is (or is perceived to be) less than full private insurance.”).

⁵ Hurricane Katrina made landfall on August 29, 2005, and the levees failed on August 30, 2005. See Joseph B. Treaster & N.R. Kleinfield, *Hurricane Katrina: The Overview; New Orleans Is Inundated As 2 Levees Fail; Much of Gulf Coast Is Crippled; Toll Rises*, N.Y. TIMES, Aug. 31, 2005 at A1, available at <http://query.nytimes.com/gst/fullpage.html?res=940CE4DF1731F932A0575BC0A9639C8B63&pagewanted=all>; Joseph B. Treaster & Kate Zernike, *Hurricane Katrina Slams into Gulf Coast; Dozens Are Dead*, N.Y. TIMES, Aug. 30, 2005, at A1, available at <http://www.nytimes.com/2005/08/30/national/30storm.html?pagewanted=all>.

⁶ See HOWARD C. KUNREUTHER & ERWANN O. MICHEL-KERJAN, AT WAR WITH THE WEATHER: MANAGING LARGE-SCALE RISK IN A NEW ERA OF CATASTROPHES 4 (2009) (explaining that “development in hazard-prone areas and increased value at risk” are key factors and climate change is “of growing concern”); Adam F. Scales, *A Nation of Policyholders: Governmental and Market Failure in Flood Insurance*, 26 MISS. C. L. REV. 3, 6 & n. 12 (2006) (describing how development has increased the cost of floods, though “global warming or cyclical climate changes may explain part of this increase”).

⁷ Since the original draft of this article was written, near-record setting water levels along the Mississippi River have exacted their toll, including the opening of spillways to flood purposefully rural areas in order to avoid catastrophic losses in larger metropolitan areas. See, e.g., Christine Hauser, *Flooding Takes Vast Economic Toll, And It’s Hardly Done*, N.Y. TIMES, May 18, 2011, at A11; Campbell Robertson, *Louisiana Spillway Opened to Relieve Flooding*, N.Y. TIMES, May 15, 2011, <http://www.nytimes.com/2011/05/15/us/15spillway.html>; A.G. Sulzberger, *As Missouri River Rises, Control Efforts Take Shape*, N.Y. TIMES, June 3, 2011, at A14. See also CHRISTINE A. KLEIN & SANDRA B. ZELLMER, MISSISSIPPI RIVER STORIES: HOW THE ROAD TO UNNATURAL DISASTER IS PAVED WITH WELL-INTENDED LAWS (forthcoming 2011), for more on the history of flooding along the Mississippi River.

would harness the strengths of the federal taxing authority⁸ to provide basic relief for flood losses occurring at an individual's primary residence.⁹ Any plan for addressing flood loss must navigate hotly debated, competing views about government intervention, redistribution, private markets, environmental protection, and property rights. This Article argues that governmental intervention in flood loss relief is inevitable, at least in the foreseeable future,¹⁰ and that the focus of that intervention should be on the *ex ante* provision of a social safety net. The program proposed in this Article is also intended to provide additional levers for addressing the complexities of flood loss, including the reduction of negative environmental externalities,¹¹ and to provide the impetus needed for harmonizing existing tax provisions and grant programs.

Part II of this Article discusses the NFIP's program for personal property¹² and outlines problems associated with the program. Overall, the

⁸ See Scott E. Harrington, *Rethinking Disaster Policy After Hurricane Katrina*, in ON RISK AND DISASTER: LESSONS FROM HURRICANE KATRINA 203, 217 (Ronald J. Daniels et al. eds., 2006) (briefly raising the possibility of a premium tax approach and stating that it is a "potentially superior approach").

⁹ The business and investment property flood losses will be addressed in a future Article.

¹⁰ See Robert H. Jerry, II & Steven E. Roberts, *Regulating the Business of Insurance: Federalism in an Age of Difficult Risk*, 41 WAKE FOREST L. REV. 835, 875 (2006) ("Although the flood insurance program has serious deficiencies, no one seriously suggests that management of this market should revert to the states."); Louis Kaplow, *An Economic Analysis of Legal Transitions*, 99 HARV. L. REV. 509, 536-50 (1986) (discussing market versus government solutions in the presence of market failure); Levmore & Logue, *supra* note 4; George L. Priest, *Government Insurances versus Market Insurance*, 28 GENEVA PAPERS ON RISK & INS. 71 (2003); Michael J. Trebilcock & Ronald J. Daniels, *Rationales and Instruments for Government Intervention in Natural Disasters*, in ON RISK AND DISASTER: LESSONS FROM HURRICANE KATRINA 89 (Ronald J. Daniels et al. eds., 2006). See, e.g., Kenneth J. Arrow, *Uncertainty and the Welfare Economics of Medical Care*, 53 AM. ECON. REV. 941 (1963) for discussion of the more general question of when government intervention is appropriate.

¹¹ See Eric J. Johnson et al., *Framing, Probability Distortions, and Insurance Decisions*, in CHOICES, VALUES, AND FRAMES 224, 231-32 (Daniel Kahneman & Amos Tversky eds., 2000) (complete shift of risk to insurer "could lead the insured to be irresponsible because he or she bears no cost of a loss").

¹² The NFIP also authorizes business coverage. See 42 U.S.C. § 4012; *Commercial Coverage: Business Property Risk*, FLOODSMART.GOV, http://www.floodsmart.gov/floodsmart/pages/commercial_coverage/business_property_risk.jsp (last updated Aug. 25, 2011, 4:18 PM) (overview of currently

NFIP fails to provide an adequate safety net as numerous individuals continue to fail to purchase flood insurance.¹³ If the NFIP were to charge actuarially fair premiums,¹⁴ the resulting increases would likely lead to even lower participation in the program among those least economically able to self-insure.¹⁵ At the same time, some individuals file repetitive loss claims, causing a significant financial drain on the program and potentially exacerbating environmental costs.¹⁶ The budget woes of the NFIP are compounded by the outsourcing of flood insurance sales and claims adjustments to private insurance companies.¹⁷ These private insurance companies charge the NFIP a flat rate for these services without having to account for actual costs.¹⁸

The NFIP's problem areas are relatively easy to enumerate, but the path to crafting a better approach is more complex. Part III discusses some of the obstacles facing any plan designed to mitigate and compensate for flood loss. Flood losses are difficult to diversify; individuals have an incentive to purchase flood insurance only for their most at-risk property; and individuals may be motivated to take less care in their decisions with

available business coverage). Discussion of NFIP business coverage as well as business-related tax provisions is outside the scope of this Article.

¹³ See, e.g., Howard Kunreuther, *Has The Time Come for Comprehensive Natural Disaster Insurance*, in ON RISK AND DISASTER: LESSONS FROM HURRICANE KATRINA 175, 175 (Ronald J. Daniels et al. eds., 2006) (For Louisiana parishes hit by Katrina, "the percentages of homeowners with flood insurance ranged from 57.7 percent . . . to 7.3 percent. . . . Only 40 percent of the residents in Orleans parish had flood insurance.").

¹⁴ See U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-10-1063T, NATIONAL FLOOD INSURANCE PROGRAM: CONTINUED ACTIONS NEEDED TO ADDRESS FINANCIAL & OPERATIONAL ISSUE, at 5-6 (2010) (finding that NFIP "is, by design, not actuarially sound").

¹⁵ See *id.* at 3 (explaining that taking steps to "make premium rates more reflective of long-term flood risks . . . would raise rates and potentially reduce participation in NFIP.").

¹⁶ See *id.* at 1 ("Only 1 percent of policies . . . account for 25 to 30 percent of claims.").

¹⁷ Before the massive flooding of 2011, the NFIP was already deeply in debt, largely because of the catastrophic losses of the 2005 hurricane season. See *id.* ("As of August 2010, NFIP's debt to Treasury stood at \$18.8 billion."). Before the 2005 hurricane season, the program had generally balanced out. See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 110-11.

¹⁸ See Aviva Abramovsky, *Insurance and the Flood*, in LAW AND RECOVERY FROM DISASTER: HURRICANE KATRINA 83, 97 (Robin Paul Malloy ed., 2009); see also U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 8-9.

respect to flood costs because of the availability of coverage. These three difficulties — known respectively as correlation,¹⁹ adverse selection,²⁰ and moral hazard²¹ — represent classic concerns in the formation of insurance markets. Part III also briefly considers possible cognitive obstacles to the provision of flood loss relief.²² For example, because flood risks are difficult to conceptualize, individuals will have problems taking the steps necessary to engage in adequate preparation, and government officials charged with aiding community preparation will be subject to the same challenges.²³

¹⁹ See DAVID A. MOSS, WHEN ALL ELSE FAILS: GOVERNMENT AS THE ULTIMATE RISK MANAGER 262 (2002) (explaining that by 1928 “[h]aving learned that individual flood risks were often highly correlated . . . insurers had apparently decided that the prospect of catastrophic flooding rendered this particular risk uninsurable”); Michelle E. Boardman, *Known Unknowns: The Illusion of Terrorism Insurance*, 93 GEO. L.J. 783, 820 (2005) (“Natural disasters are highly correlated, and difficult to ‘uncorrelate’ because those who are not at high risk do not seek to transfer their risk.”); see also *infra* Part III.A for discussion regarding why even national, private insurance companies face correlation difficulties with respect to flood loss.

²⁰ See Tom Baker, *Containing the Promise of Insurance Adverse Selection and Risk Classification*, 9 CONN. INS. L.J. 371, 378 (2003) (arguing that “risk classification itself can create a kind of adverse selection” since insurers may “select risks in a manner that is *adverse* to the insurance pool”); Kaplow, *supra* note 10, at 543-44 (explaining that pricing to cover high-risk individuals will cause lower-risk individuals to drop out, which will cause insurance companies to increase rates again and so motivate even more lower-risk individuals to drop coverage, and so on until it is possible that “no insurance would be offered”).

²¹ See Tom Baker, *On the Genealogy of Moral Hazard*, 75 TEX. L. REV. 237, 239 (1996) (explaining that in economic literature the term “refers to the tendency for insurance against loss to reduce incentives to prevent or minimize the cost of loss.”); Kaplow, *supra* note 10, at 537 (with insurance “actors have less incentive to avoid” losses); Kunreuther, *supra* note 13, at 183 (“[D]isaster assistance is purported to create a type of Samaritan’s dilemma: providing assistance after a catastrophe reduces the economic incentives of potential victims to invest in protective measures prior to a disaster.”).

²² See *infra* Part III.B.

²³ See Johnson et al., *supra* note 11, at 225 (“A rational, risk-neutral consumer would purchase coverage at an actuarially fair price that is equivalent to the expected loss. . . . In practice, the story is apparently not that simple.”); Kaplow, *supra* note 10, at 548 (stating that the “strongest case for some government response to risk is presented by situations in which certain actors underestimate the likelihood of loss”).

Part IV argues that utilizing tax system components may provide a strong course for meeting the complexities of flood loss coverage and mitigation, though it also discusses the challenges that would face such an approach. Additionally, Part IV presents an outline of such a tax-system infused flood loss security program. The proposed program would be administered jointly by the Treasury (IRS) and Homeland Security (FEMA) and would mandate minimum coverage for all individuals as to the contents of their primary residences.²⁴ Coverage for a home's structure would also be mandatory but should be designed to limit repetitive loss claims. Rewards as well as penalties could be built into the system in order to better manage flood preparation and community participation. For example, the proposed flood security plan could charge rates that allow for tax refunds in the case of good results — e.g., no claim filed in a particular year.²⁵ Income tax refunds appear to be highly satisfying given the amount of over-withholding that occurs in the income tax system.²⁶

Part V explores the current patchwork of tax rules as they relate to post-disaster assistance, pre-disaster flood mitigation grant programs, and insurance payouts. Part V also recommends steps for harmonizing these rules with the proposed flood loss security program. Part VI is a brief conclusion.

II. THE NATIONAL FLOOD INSURANCE PROGRAM

The National Flood Insurance Program (NFIP)²⁷ is administered by the Federal Emergency Management Agency (FEMA),²⁸ which is a part of the Department of Homeland Security. The NFIP has roots dating back to the early 1950s²⁹ and the early legislation introduced structural components

²⁴ Mandates have long been recognized as a solution to the adverse selection problem. *See infra* Part III.A. If such a mandate is, however, politically unpalatable, coverage could be mandatory for high and moderate risk residences while opt-out coverage could be available for lower-risk residences. *See infra* Part IV.B.

²⁵ *See* Johnson et al., *supra* note 11, at 232-33, 238 (describing insureds' preference for rebates over deductibles).

²⁶ *See* Lee Anne Fennell, *Hyperopia in Public Finance*, in BEHAVIORAL PUBLIC FINANCE 141, 148-52 (Edward J. McCaffery & Joel Slemrod eds., 2006).

²⁷ 42 U.S.C. §§ 4011-31 (2006).

²⁸ *See* 42 U.S.C. § 4011(a).

²⁹ *See* HOWARD KUNREUTHER & DOUGLAS C. DACY, THE ECONOMICS OF NATURAL DISASTERS 259 (1969), for more on the history behind the NFIP; MOSS, *supra* note 19, at 262-63; Abramovsky, *supra* note 18, at 92; David A. Grossman,

that, while well intentioned, contribute to the weakness of the NFIP today. This Part provides an overview of the current state of the program.

A. COMMUNITY PARTICIPATION

Early flood insurance legislation attempted to motivate communities to take flood mitigation steps by tying the availability of insurance coverage to community adherence to floodplain management regulations.³⁰ Even today, individuals are not able to participate in the NFIP unless their communities agree to abide by various regulations intended to mitigate flood loss.³¹ As to communities who fail to participate, federal grants, disaster relief, and federal mortgage insurance are “unavailable for the acquisition or construction of structures located or to be located” in high-risk areas.³² Currently, over twenty thousand

Flood Insurance: Can a Feasible Program be Created?, 34 LAND ECON. 352 (1958).

³⁰ See 42 U.S.C. §§ 4022, 4102 (2006) (community participation requirements). The regulatory requirements are extensive. See 44 C.F.R. pts. 60, 64; FEMA, *Floodplain Management Requirements: A Study Guide and Desk Reference for Local Officials*, FEMA (Aug. 11, 2010), http://www.fema.gov/vplan/prevent/floodplainfm_sg.shtm; Abramovsky, *supra* note 18, at 92.

³¹ 42 U.S.C. § 4012(c) (2006); see Edward T. Pasterick, *The National Flood Insurance Program*, in PAYING THE PRICE: THE STATUS AND ROLE OF INSURANCE AGAINST NATURAL DISASTERS IN THE UNITED STATES 125, 131 (1998) (Howard Kunreuther & Richard J. Roth, Sr., eds.) (discussing responsibility of local community in “adopting and enforcing these floodplain management standards”). Relatively few individuals would be affected by the non-participation of the local community because “[m]ost flood-prone communities that have elected not to participate are communities whose areas of serious flood risk are either very small or have few if any structures.” *Id.* at 129; FEMA, NATIONAL FLOOD INSURANCE PROGRAM: MANDATORY PURCHASE OF FLOOD INSURANCE GUIDELINES 2 (2007), available at www.fema.gov/library/viewRecord.do?id=2954 (“If a community does not participate in the program, property owners in that jurisdiction are not able to purchase federally backed flood insurance.”). Individuals living in non-participating communities would have to rely on post-flood government assistance or on the virtually nonexistent private flood insurance market. Abramovsky, *supra* note 18, at 126 (“[P]rivate insurers do write limited amounts of flood coverage, usually for commercial insureds”).

³² See 42 U.S.C. § 4106 (2006); see also FEMA, *supra* note 31, at 2. A 1968 Act did contain a short-lived penalty at the individual level that had community participation implications: if the individual’s community participated and the individual failed to purchase flood insurance coverage after one year, then such

communities participate.³³ Since 1990 communities have also been able to elect to comply with stronger standards through the Community Rating System.³⁴ Participation in the Community Rating System program yields credits that have the effect of reducing flood insurance premiums throughout the community.³⁵ Currently, nearly twelve hundred communities participate in the Community Rating System program, which while representing only 5 percent of all NFIP communities includes approximately 67 percent of NFIP policyholders.³⁶ In spite of widespread community participation, individual residents will not necessarily have flood insurance because, as will be discussed more fully in the next section, purchase of coverage is largely optional.³⁷

Participation by a community in the NFIP does not, of course, ensure that a local community is actually compliant.³⁸ FEMA must determine whether local building codes and permitting processes on their face adhere to the federal guidelines and must also examine whether communities actually follow facially adequate ordinances.³⁹ Communities may further complicate FEMA's job by pushing back against guidelines

individuals were to be denied post-flood federal assistance. 42 U.S.C. § 4021 (repealed); *see also* Abramovsky, *supra* note 18, at 92-93.

³³ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 4.

³⁴ 42 U.S.C. § 4022(b); *see also* Pasterick, *supra* note 31, at 135-36 (describing system).

³⁵ *See* Pasterick, *supra* note 31, at 135. Credits are based on "estimated reduction in flood and erosion damage risks resulting from the measures adopted by the community under the program." 42 U.S.C. § 4022(b)(3).

³⁶ Email from William L. Trakimas, Director of Natural Hazards (Sept. 8, 2011) (on file with authors) ("Currently 1192 communities participate nationwide . . . receiv[ing] a discount which is about \$292M annually."). In 1998, roughly 900 communities participated, which similarly represented 5 percent of NFIP communities but included over 63 percent of NFIP policyholders. Pasterick, *supra* note 31, at 137.

³⁷ *See infra* Part II.B.

³⁸ *See, e.g.,* KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 17 (noting that "25 percent of the insured losses from Hurricane Andrew in 1992 could have been prevented through better building code compliance and enforcement"); *see also* Raymond J. Burby, *Hurricane Katrina and the Paradoxes of Government Disaster Policy: Bringing About Wise Governmental Decisions for Hazardous Areas*, 604 ANNALS AM. ACAD. OF POL. & SOC. SCI. 171, 178 (2006) (describing how many local governments fail to enforce the minimum building requirements need to participate in the NFIP).

³⁹ *See* Pasterick, *supra* note 31, at 131.

whose implementation they perceive to be too costly.⁴⁰ New floodplain management regulations often contain transition rules or grandfather provisions,⁴¹ possibly in order to minimize political fallout. The political dimensions of putting a community on probation or pulling NFIP eligibility⁴² may also constrain enforcement.⁴³

Even assuming full compliance with floodplain regulations, the regulations, in conjunction with other flood loss mitigation programs, may have unintended consequences. Individuals may be overly confident in the ability of federal, state, and local authorities to manage flood loss through artificial containment and diversion projects and thus increase the direct and externalized costs of floods. That is, development may increase in areas that have been rendered “safe” through community planning.⁴⁴ (Alternatively, development may occur first under the assumption that with

⁴⁰ See Peter G. Gosselin, *On Their Own in Battered New Orleans*, in ON RISK AND DISASTER: LESSONS FROM HURRICANE KATRINA 15, 22-23 (Ronald J. Daniels et al., eds., 2006) (describing among New Orleans residents that regulation changes would make it difficult to maintain flood insurance eligibility); see also DENNIS C. MUELLER, PUBLIC CHOICE III 343-47, 473 (2003) (describing formation of interest groups and agency capture).

⁴¹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 14.

⁴² See 42 U.S.C. § 4023 (2006) (disallowing flood insurance coverage for communities violating state and local land use law); 44 C.F.R. § 59.24 (2010) (suspension regulations); 44 C.F.R. § 61.16 (2010) (additional premium charged in communities on probation); 44 C.F.R. pt. 73 (2010) (violations of state and local zoning law); see also Pasterick, *supra* note 31, at 131 (describing probation and suspension process).

⁴³ See Pasterick, *supra* note 31, at 131 (“[T]here has never been a comprehensive assessment of the level of compliance nationwide or of the overall effect of program standards on local development patterns.”).

⁴⁴ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 263 (discussing how government actions may make residents feel safe when in fact they remain vulnerable); Burby, *supra* note 38, at 176 (federal policy in New Orleans contributed “directly to the devastation of Hurricane Katrina” by encouraging development in hazardous areas and diverting resources away from areas that could have benefitted from improvements); Klein & Zellmer, *supra* note 4, at 1518 (describing the “foolhardiness of . . . attempting to keep the water away from the people through artificial flood control”); Scales, *supra* note 6, at 6 (discussing how “[f]lood control projects merely buy time” but also attract “[r]esidential and commercial development . . . often resting on long-term assumptions about the suitability of the area for development”).

increased development, loss mitigation will be undertaken.⁴⁵) Individual homeowners and renters may then rely not only on visible governmental mitigation efforts but may be further reassured by the presence of developers. If, however, the safety measures fail (or fail to materialize) the flood costs will be even higher because of the increased development.⁴⁶ The failure of the levees in New Orleans is among the most vivid examples of the risk of relying on manmade structures to turn back nature.⁴⁷ Although individuals residing in New Orleans had the option to purchase flood insurance, the majority of residents did not do so and were not required to do so⁴⁸ (the same would almost certainly hold true in any U.S. community⁴⁹). Individuals may well not have understood that risk was still present in spite of (or because of) the levees.⁵⁰

⁴⁵ Finn E. Kydland & Edward C. Prescott, *Rules Rather Than Discretion: The Inconsistency of Optimal Plans*, 85 J. POLITICAL ECON. 473, 477 (1977) (“[T]he rational agent knows that, if he and others build houses there [in the flood plain], the government will take the necessary flood-control measures. Consequently, in the absence of a law prohibiting the construction of houses in the flood plain, houses are built there, and the army corps of engineers subsequently builds the dams and levees.”); see also KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 262 (describing Nobel Prize-winning work of Kydland and Prescott, including flood plain example showing “that a discretionary policy, which may be optimal given the current situation, may not necessarily result in a socially optimal policy in the longer run”).

⁴⁶ See generally Klein & Zellmer, *supra* note 4; Scales, *supra* note 6, at 13 (“[F]loodplain management (rather than floodplain abandonment) encouraged development and, thus, concentrated rather than dispersed economic risks of flooding.”).

⁴⁷ The 2011 flooding along the Mississippi river is also illustrative of this lesson. See Editorial, *A New Flood, Some Old Truths: The Mississippi Tells Us, Again, To Change The Way We Manage Water*, N.Y. TIMES, May 28, 2011, at A22 (“Years of mismanagement of the vast Mississippi River ecosystem—the relentless and often inadvisable construction of levees and navigation channels, the paving over of wetlands, the commercial development of flood plains . . . have made the damage worse than it might otherwise have been. . . . Nobody ever beats the river.”).

⁴⁸ See Jerry & Roberts, *supra* note 10, at 877 (“[T]he percentage of homes with flood insurance policies in coastal parishes of Louisiana affected by Hurricane Katrina ranged from 7% in St. James Parish to 57.7% in St. Bernard Parish, with only 40% of homes in Orleans Parish having this coverage.”); Scales, *supra* note 6, at 15 (“[F]ewer than one-in-ten residents along the Gulf Coast of Mississippi are believed to have held flood insurance prior to Katrina.”).

⁴⁹ See LLOYD DIXON ET AL., THE NATIONAL FLOOD INSURANCE PROGRAM’S MARKET PENETRATION RATE: ESTIMATES AND POLICY IMPLICATIONS xvi (2006)

While flood mitigation programs have unintended consequences, halting mitigation programs is likely to be even more problematic. First, mitigation does work⁵¹ albeit only up to a point — though often an unknown point at that. Second, outright prohibitions on development by the federal government are problematic,⁵² and once development has occurred, and if the potential disaster is big enough, the federal government will find it politically untenable to fail to provide any mitigation.⁵³ Even assuming developers understand the riskiness of their building projects, they may be able to shift the flood risk to the ultimate owners and tenants,⁵⁴ who are sure to elicit (and likely to deserve) a more sympathetic response than the original developers. Thus, continuance of flood mitigation programs, including community participation, appears to be an uneasy necessity, though steps could clearly be taken to use mitigation more judiciously and development prohibitions less sparingly.⁵⁵ As will be discussed in Part IV, even though this Article does not directly address the role of developers

(“Even though approximately one-third of NFIP policies are written outside SFHAs [high-risk areas], the market penetration rate outside SFHAs is only about 1 percent.”).

⁵⁰ See *infra* Part III.B (discussing possible reasons, including cognitive shortcuts and biases, for low participation in flood insurance).

⁵¹ See Pasterick, *supra* note 31, at 131-32 (discussing how flood plain regulations have, at least in the Midwest, “discourage[d] floodplain development through the increased costs in meeting floodplain management requirements and the cost of an annual flood insurance premium”); David Welky, *When the Levee Doesn’t Break*, N.Y. TIMES, May 11, 2011, at A25 (arguing that “the extent of the [2011 Mississippi flood] damage probably won’t come close to the losses of life and property seen in the historic flood of January 1937. . . .—proof that after nearly 75 years, the federal government has finally gained the upper hand on a river system once thought uncontrollable.”).

⁵² See Pasterick, *supra* note 31, at 131 (noting rejection by NFIP of federal override of local regulation because the NFIP “has consistently taken the position that federal land use regulation at the local level is illegal, and, in any case, would be unworkable”).

⁵³ See Kydland & Prescott, *supra* note 45, at 477 (theorizing that the “the rational agent knows that, if he and others build houses there [in the flood plain], the government will take the necessary flood-control measures”).

⁵⁴ Cf. Pasterick, *supra* note 31, at 131-32 (discussing report in Midwest suggesting that “[d]evelopers have the added incentive of wanting to avoid marketing flood-prone property.”).

⁵⁵ See *id.* at 154 (noting “vital connection between the availability of flood insurance and the local community enforcement of floodplain management provisions”).

and other commercial enterprises, integrating residential flood loss coverage with the tax system could provide an opportunity to craft additional levers for balancing social safety net concerns with constraints on unwise development.

B. INDIVIDUAL PARTICIPATION

Individuals are required to purchase flood insurance only in a limited set of circumstances. Regulated lending institutions,⁵⁶ government-sponsored enterprises for housing (e.g., Fannie Mae and Freddie Mac), and federal agency lenders⁵⁷ must require flood insurance as a condition to closing on loans secured by property in high-risk flood zones.⁵⁸ “High-risk” indicates that there is a 1% or greater chance of a flood in a particular year⁵⁹ — that is, the property lies within the one-hundred year flood plain.

⁵⁶ 42 U.S.C. § 4003(a)(10) (2006) (includes “any bank, savings and loan association, credit union, farm credit bank, Federal land bank association, production credit association, or similar institution subject to the supervision on a Federal entity for lending regulation”). The statute directs the federal entities for lending regulation to promulgate regulations applicable to these institutions. These federal entities have adopted such regulations. *See* 12 C.F.R. pt. 22 (2010) (by Office of the Comptroller of the Currency); 12 C.F.R. § 208.25 (2010) (Federal Reserve System); 12 C.F.R. pt. 339 (2010) (Federal Deposit Insurance Corporation); 12 C.F.R. pt. 572 (2010) (Office of Thrift Supervision); 12 C.F.R. § 614.4920 (2010) (Farm Credit Administration); 12 C.F.R. pt. 760 (2010) (National Credit Union Administration).

⁵⁷ 42 U.S.C. § 4003(a)(7) (2006) (defining these agencies as “Federal agenc[ies] that makes direct loans secured by improved real estate or a mobile home”). *See* FEMA, *supra* note 31, at 26 (entities include Federal Housing Administration, Small Business Administration, Department of Veterans Affairs, and the U.S. Department of Agriculture).

⁵⁸ 42 U.S.C. § 4012a(b) (2006) (lender mandate); 42 U.S.C. § 4104a (2006) (notice requirements). *See* FEMA, *supra* note 31, at 2-4 (the only lenders and services excluded are those “who are not federally regulated and that do not sell loans to . . . Fannie Mae . . . Freddie Mac,” or other government-sponsored entities.).

⁵⁹ FEMA literature often uses the term “special flood hazard area” but “High-risk flood areas” and “special flood hazard areas” are synonymous. *Compare* U.S. Gov’t Accountability Office, *supra* note 14, at 14 *with* FEMA, *supra* note 31, at GLS 9.

This is also called the 100-year flood plain. *See* 44 C.F.R. § 59.1 (2010) (defining “100-year flood” as “the flood having a one percent chance of being equaled or exceeded in any given year”); FEMA, *supra* note 31, at GLS 9. But

All mapped areas with lower than 1% chance per year of flooding are in low or moderate-risk zones;⁶⁰ yet such zones historically lead to about 25 percent of NFIP claims.⁶¹ Since relatively few individuals purchase insurance if they reside outside a high-risk zone, such policies constitute such a significant portion of NFIP claims suggests that the 1% benchmark is problematic.⁶²

The lender mandate does not apply to properties outside of high-risk flood zones. The requirement also does not apply to properties located in non-participating communities since individuals in those areas are not eligible to purchase flood insurance.⁶³ Under the most recent changes to the

such terminology can mislead individuals into thinking that a flood will only occur once in a hundred years and is downplayed (or eliminated) in public education information. See Pasterick, *supra* note 1, at 130 (“The term ‘100-year flood’ is problematic for the NFIP. It is a term of convenience intended to convey probability but has had the adverse effect of giving floodplain residents, who tend to interpret it in chronological terms, a false sense of security.”).

FEMA has attempted to help people understand the risk assessments by anchoring this to a more readily understood marker: the 30-year mortgage. Thus, its public education website explains that high-risk “equates to a 26% chance of flooding over the life of a 30-year mortgage.” Nat’l Flood Ins. Program, FLOODSMART.GOV, THE OFFICIAL SITE OF THE NFIP, (last visited Aug. 25, 2011, 4:17 PM), <http://www.floodsmart.gov/floodsmart/>.

⁶⁰ Nat’l Flood Ins. Program, FLOODSMART.GOV, THE OFFICIAL SITE OF THE NFIP, (last visited Aug. 25, 2011, 4:17 PM), <http://www.floodsmart.gov/floodsmart>.

⁶¹ FEMA, *supra* note 31, at 5.

⁶² This estimate may be too low. See Burby, *supra* note 38, at 177 (stating that “most flood losses in the United States stem from less frequent flood events” and citing studies suggesting a range of 66% to 83% of losses arising from areas outside the one-hundred-year flood zone). The Association of State Floodplain Managers has recommended that a five-hundred-year flood plain be used as the better benchmark for levees. ASSOCIATION OF STATE FLOODPLAIN MANAGERS, NATIONAL FLOOD POLICY CHALLENGES: LEVEES: THE DOUBLE-EDGED SWORD 3-5 (2007), available at http://www.floods.org/PDF/ASFPM_Levee_Policy_Challenges_White_Paper.pdf. See also Burby, *supra* note 38, at 177 (discussing proposal by Association of State Floodplain Managers).

⁶³ FEMA, *supra* note 31, at 5. In the case of a non-participating community, “a lender is still required to inspect any flood maps to determine flood hazard risk and provide notice of such risk.” *Id.* at 2. See 42 U.S.C. 4106(b) (2006) (requiring regulations on notice). Prior to 1977, regulated lending was prohibited in communities that did not participate. The change was implemented by statute. Housing and Community Development Act of 1977, Pub. L. 95-128 § 703(a), 91 Stat. 1144. See also FEMA, *supra* note 31, app. at 1-3.

NFIP in 2004, lender-mandated insurance must remain in force over the life of the loan⁶⁴ and must be monitored by loan servicers for loans sold to Fannie Mae and Freddie Mac.⁶⁵ Various specific rules have been enacted to facilitated compliance.⁶⁶ For example, if the loan requires an escrow—for example, for real property taxes or homeowner’s insurance — flood insurance premiums are also required to be escrowed.⁶⁷

FEMA has no statutory authority to enforce this lender mandate,⁶⁸ instead, each agency with direct oversight over the covered lender is to enforce the requirement.⁶⁹ A 2006 study done by RAND estimated national compliance with the mandate at 75-80 percent, but with significant variation across regions.⁷⁰ Given the recent turmoil in the lending and housing market, including problems with administrative agency oversight and complicated securitization structures, it seems fair to wonder about the extent to which these lender flood insurance mandates have been working in recent years.⁷¹ For high-risk properties not covered by the lender

⁶⁴ FEMA, *supra* note 31, at 5.

⁶⁵ *Id.* at 25.

⁶⁶ *See generally id.* at 23-60.

⁶⁷ 42 U.S.C. § 4012a(d) (2006). *See* FEMA, *supra* note 31, at 39 (discussing requirement).

⁶⁸ FEMA, *supra* note 31, at vii. *See* 42 U.S.C. § 4012a (2006).

⁶⁹ *See supra* note 56 (citing regulatory provisions and listing these agencies); FEMA, *supra* note 31, at 59-60 (Civil penalties may be assessed. “As of November 30, 2006, a total of 119 banks had been assessed nearly \$1.3 million in penalties, for various violations of the 1994 Reform Act.” Regulators may also impose other sanctions including “unsatisfactory bank ratings, memoranda of understanding, and, ultimately, cease and desist orders.” Private individuals, including borrowers, have no cause of action against lenders who have failed to enforce the mandate.).

⁷⁰ DIXON ET AL., *supra* note 49. *See* Howard Kunreuther & Mark Pauly, *Rules Rather Than Discretion: Lessons from Hurricane Katrina*, 33 J. RISK UNCERTAINTY 101, 107 (2006) (discussing evidence that suggesting “that some banks, which were expected to enforce the requirements that individuals in high-hazard areas purchase flood coverage, looked the other way.”); Scales, *supra* note 6, at 14-15 (discussing RAND study and other scholarship on takeup rates). The failure of lenders independently to require flood insurance is a mystery, particularly given their insistence on general casualty insurance. *See also* Scales, *supra* note 6, at 17-19 (discussing possible theories for lender behavior with respect to flood insurance).

⁷¹ *See, e.g.*, Binyamin Appelbaum, *Without Loan Giants, 30-Year Mortgage May Fade Away*, N.Y. TIMES, Mar. 3, 2011, http://www.nytimes.com/2011/03/04/business/04housing.html?_r=1&scp=1&sq=Without%20Loan%20giants,%2030-year%20mortgage%20may%20fade%20away&st=cse (discussing

mandate, the same RAND study estimated approximately a 50% take-up rate.⁷²

In addition to the lender mandate, the NFIP has only one additional means of applying legal pressure on an individual's decision to purchase coverage. Under the current NFIP, individuals may receive government assistance after a disaster even if they were eligible for, but failed, to purchase flood insurance, but a condition of the assistance is that the individual purchase flood insurance in the future. Failure to purchase insurance then can be used to withhold assistance if flood loss help again becomes necessary.⁷³ Whether this penalty is actively enforced is another question,⁷⁴ particularly in the immediate aftermath of high-impact events.⁷⁵

The NFIP has no ability to deny coverage if individuals are eligible to purchase the insurance.⁷⁶ As a result of repetitive losses, the GAO

possible demise of Fannie Mae & Freddie Mac and resulting changes to the housing market).

⁷² DIXON ET AL., *supra* note 49, at xvi.

⁷³ 42 U.S.C. § 5154(b); *see also* FEMA, *supra* note 31, at 7 (discussing requirement); Pasterick, *supra* note 31, at 153 (discussing history of this requirement and noting it “has its greatest potential impact on grant recipients, who are generally in lower-income categories than those receiving loans and thus less likely to be able to afford insurance. Whether the threat of denial of future federal assistance will have the intended effect of promoting insurance purchase among this segment of the population remains to be seen.”).

⁷⁴ *See* Scales, *supra* note 6, at 13 (“[T]he NFIP’s enforcement mechanisms are limited and not credibly invoked.”).

⁷⁵ *See* Levmore & Logue, *supra* note 4, at 292-93 n.82 (predicting “that public sympathy and interest-group pressure would make enforcement of that restrictive very difficult”).

⁷⁶ U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 14, at 1. *Contra* Scales, *supra* note 6, at 33-34 (stating the NFIP does however, rigidly deny claims filed more than 60 days after a loss, even though the difficulties involved in a flood make filing the Paperwork difficult – perhaps especially for less sophisticated individuals). *But see* 16 U.S.C. § 3503 (2006) (establishing these systems); 42 U.S.C. § 4028 (stating the NFIP is not available in certain zones designated as with the Coastal Barrier Resources System); Emergency Management and Assistance 44 C.F.R. §§ 71.1, 71.3 (2010) (implementing regulations); Pasterick, *supra* note 31, at 146-47 (discussing history of legislation); *id.* at 146 (stating the Legislation applies primarily to zones within barrier islands); *id.* at 146 (stating communities may have some areas within such zones and others outside, and “[c]onsistent enforcement . . . is difficult . . . [and] the NFIP must depend on the vigilance of insurance agents to distinguish which areas of a community are eligible for coverage and which are not.”) (alteration in the original); *id.* at 146-47 (“A review conducted in 1992 by the General Accounting [sic] Office (GAO) found not only

estimates one percent of policies “account for 25 to 30 percent of claims.”⁷⁷ The dollar amounts associated with repetitive loss claims are, of course, only part of the true cost of such claims since frequently such properties are built in environmentally fragile locations.⁷⁸

Although the NFIP covers a relatively low number of individuals, the 2005 hurricane season’s demands on the NFIP were staggering and overwhelmed the NFIP. FEMA had to invoke its authority to borrow funds from the U.S. Treasury and seek additional appropriations.⁷⁹ As of August 2010, FEMA’s debt stood at \$18.8 billion;⁸⁰ it remains unlikely that the program will be able to repay this amount.⁸¹ The billions in payouts made under the NFIP are still small, however, in comparison to the total cost to the government of the disaster.⁸²

C. COVERAGE LIMITS, FLOOD MAPS, AND RATES

The maximum coverage currently available under the NFIP is \$100,000 for personal property and \$250,000 for residential real estate.⁸³ The premium rate structure varies with coverage, deductible, and, most importantly, the risks associated with the property to be insured.⁸⁴ The highest sample premium (\$5,903) listed on FEMA’s website is for a coastal area, high-risk residence and contents insured for the full available coverage with a \$2,000 deductible.⁸⁵ Individuals may purchase coverage

that significant new development continued to occur in certain CBRS units after the law was enacted, but also that NFIP coverage was written on 9 percent of the residences in the units sampled.”).

⁷⁷ U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 14, at 1.

⁷⁸ See Klein & Zellmer, *supra* note 4, at 1508-10 (discussing the “value of healthy wetlands”).

⁷⁹ 42 U.S.C. §§ 4016, 4017(b)(1), (b)(3), 4127; see Burby, *supra* note 38, at 177 (discussing past history of operating losses and use of this authority); Pasterick, *supra* note 31, at 138-39 (discussing the same).

⁸⁰ U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 14, at 5, 14.

⁸¹ *Id.* at 5.

⁸² See Jerry & Roberts, *supra* note 10, at 876-77 (“[T]otal government expenditures could eventually exceed \$200 billion.”).

⁸³ 42 U.S.C. §§ 4013-4015 (statutory authorization for setting various coverage terms and rates); see also *Residential Coverage Policy Rates*, FLOODSMART.GOV (Jan. 1, 2011), http://www/floodsmart.gov/floodsmart/pages/residential_coverage/policy-rates.jsp.

⁸⁴ See Emergency Management and Assistance, 44 C.F.R. pt. 61 (2010) (coverage, rates & deductibles).

⁸⁵ See *Residential Coverage Policy Rates*, *supra* note 83.

only for residences and their contents.⁸⁶ Thus, cars are not covered,⁸⁷ but there is no limit to the number of residences for which an individual may purchase flood insurance.⁸⁸ Special restrictions do apply to basements and lower-level crawlspaces.⁸⁹ Further, “flood” under the NFIP generally does not cover subsidence⁹⁰ (which, incidentally, leaves a gap in coverage availability since private insurers also generally exclude subsidence⁹¹).

Although individuals under-purchase flood insurance, possibly because of perceptions that the rates are too high,⁹² in fact even the full risk rates charged are not actuarially sound.⁹³ FEMA is charged with maintaining flood risk maps,⁹⁴ but such mapping is difficult given the contingencies that must be modeled and the costs involved in generating accurate assessments. Maps cannot remain static since flood risks will change over time both through natural occurrences and manmade development. Many FEMA maps are badly in need of updating and also often fail to take into account important risks.⁹⁵

In addition to any scientific or budgetary difficulties surrounding the creation of accurate flood maps, after updates, if FEMA changes maps,

⁸⁶ See 44 C.F.R. pt. 61 App. A (2)-(3) (stating that renters insurance is available as well as condo insurance).

⁸⁷ *Residential Coverage: What's Covered*, FLOODSMART.GOV, http://www.floodsmart.gov/floodsmart/pages/residential_coverage/whats_covered.jsp (last visited Aug. 25, 2011).

⁸⁸ *Residential Coverage: Policy Rates*, *supra* note 83 (“Single-family dwellings that are primary residences and insured to the maximum amount of insurance available under the program or no less than 80% of the replacement cost at the time of may qualify for replacement cost claim settlement. All other buildings and contents will be adjusted based on their Actual Cash Value (depreciated cost).”).

⁸⁹ *Residential Coverage: What's Covered*, *supra* note 87.

⁹⁰ *Contra* 44 C.F.R. pt. 61 App. A(1) § II(A) (Coverage is, however, available for “subsidence of land along the shore of a lake or similar body of water as a result of erosion or undermining caused by waves or currents of water exceeding anticipated cyclical levels that result in a flood . . .”).

⁹¹ See Scales, *supra* note 6, at 35.

⁹² See *infra* Part III.B.

⁹³ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 5-6.

⁹⁴ 42 U.S.C. §§ 4101(a), (e)-(i) (2006) (requiring establishment and publication of information about flood risk zones); see also 44 C.F.R. § 64.3 (description of flood insurance maps); 44 C.F.R. pt. 65 (special hazard mapping).

⁹⁵ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 7; see also Pasterick, *supra* note 31, at 144-46 (describing problem of erosion in general).

FEMA will often be viewed as the proverbial bearer of bad news.⁹⁶ As discussed above, rate increases or more stringent floodplain management requirements may have political repercussions,⁹⁷ and FEMA has generally adopted the administrative practice of grandfathering in current policyholders to the prior rate.⁹⁸ In addition to administratively crafted grandfathering rules, subsidized rates are required by statute to apply to policyholders who own “structures that were built before floodplain management regulations were established.”⁹⁹ These structures date to the origination of the NFIP, and even forty-plus years later, nearly 25 percent of NFIP policies receive these subsidized rates.¹⁰⁰ These properties also “experience as much as five times more flood damage than compliant new structures that are charged full-risk rates.”¹⁰¹

D. OUTSOURCING AND THE NFIP

The federal government sets the flood insurance terms and bears all of the risks associated with the program, marketing, sales, yet claims adjustments are increasingly handled by private insurers through the “Write Your Own” (WYO) Program.¹⁰² Under the program, for example, a policyholder could buy flood insurance from Allstate although the actual product is only available through the NFIP.¹⁰³

Utilization of private insurance companies to participate in the flood insurance program may have been intended to help market the

⁹⁶ NFIP statute and regulations require consultation with local officials and the regulations provide various procedures for appealing flood elevation and other flood map determinations. 42 U.S.C. § 4107; 44 C.F.R. pt.66 (consultation with local officials); 44 C.F.R. pt. 67 (flood elevation determination appeals); 44 C.F.R. pt. 68 (administrative hearing procedures); 44 C.F.R. pt. 70 (procedures for map correction); 44 C.F.R. pt. 72 (procedures and fees for processing map changes).

⁹⁷ See *supra* Part II.A.

⁹⁸ See *supra* Part II.A.

⁹⁹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 6; see 42 U.S.C. § 4015.

¹⁰⁰ *Id.* at 5-6; see also Pasterick, *supra* note 31, at 132-34 (describing subsidized rates applicable to pre-flood-insurance-rate-map structures); Scales, *supra* note 6, at 16 (“As of this writing, 38 years have passed, and approximately 28% [in 2006] of NFIP policies remain subsidized. This in fact reflects substantial progress, as the subsidization rate was originally 70%.”) (alteration in the original).

¹⁰¹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 5-6.

¹⁰² 42 U.S.C. § 4081; 44 C.F.R. § 62.23; see Abramovsky, *supra* note 18, at 96 (describing WYO program); Scales, *supra* note 6, at 14 (describing the same).

¹⁰³ Abramovsky, *supra* note 18, at 96.

program and provide better information to individuals regarding their financial alternatives.¹⁰⁴ WYO policies have increased dramatically as a percentage of flood insurance purchases.¹⁰⁵ By September 2008, ninety WYO insurance companies administered almost ninety-seven percent of approximately 5.6 million policies in force.¹⁰⁶ By comparison, in 1986, forty-eight WYO companies handled just under half of all policies.¹⁰⁷ While WYO policies may be a high percentage of the total outstanding policies, it is not clear whether the WYO has indeed helped increase total participation since participation in the NFIP remains low.¹⁰⁸

In creating the WYO program, the federal government may also have been seeking to lower its administrative costs.¹⁰⁹ But if so, the program is flawed. The WYO companies are paid a flat rate and are not required to account for actual costs incurred.¹¹⁰ The U.S. Government Accountability Office (GAO) studied the difference between the fee received and actual costs for six WYO insurers from 2005 through 2007 and found “that the payments exceeded actual expenses by \$327.1 million, or 16.5 percent of total payments made.”¹¹¹ The GAO has also determined that WYO insurers “did not strategically market the product” in spite of a

¹⁰⁴ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 85 (explaining that WYO was supposed to be a win-win allowing the NFIP to benefit from marketing by private insurance); Scales, *supra* note 6, at 14 (“The WYO program seemed an ideal way to remedy the NFIP’s persistent failure to sell many flood policies.”).

¹⁰⁵ Abramovsky, *supra* note 18, at 97.

¹⁰⁶ GOVERNMENT ACCOUNTABILITY OFFICE, GAO-09-455, FLOOD INSURANCE: OPPORTUNITIES EXIST TO IMPROVE OVERSIGHT OF THE WYO PROGRAM 3 (2009) (Report to the Ranking Member, Committee on Banking, Housing, and Urban Affairs, U.S. Senate), *available at* <http://www.gao.gov/new.items/d09455.pdf>.

¹⁰⁷ *Id.* at 3; *see also* Abramovsky, *supra* note 18, at 97.

¹⁰⁸ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 85 (“Despite this potentially synergistic effort between the NFIP and private companies, take-up rates for flood insurance have historically been low.”); Scales, *supra* note 6, at 14-15 (discussing participation rates and stating that “the inception of the WYO program had a very modest impact on flood insurance participation”).

¹⁰⁹ See U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 14, at 3-4.

¹¹⁰ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 83 (“More than thirty percent of each dollar paid for flood insurance coverage goes to private insurers Over the period of 1968 to 2005, these private insurers received over \$7.4 billion (excluding the loss adjustment expenses for which we do not have data) in fees.”); Abramovsky, *supra* note 18, at 97.

¹¹¹ U.S. GOV’T ACCOUNTABILITY OFFICE, *supra* note 14, at 9.

bonus structure that was added to the standard flat-rate compensation system.¹¹²

In addition to the problems that arise in having WYO insurers market both their own policies and government policies, WYO will also act as the adjusters for both their private policies and the government policies in the aftermath of a disaster.¹¹³ Thus, the same insurer will be deciding whether to categorize damage as flood damage (covered by the NFIP) or as wind damage (covered by private insurance).¹¹⁴ In the aftermath of Hurricane Katrina, press accounts reported that the WYO companies boosted flood claims in order to minimize wind damage payouts.¹¹⁵

III. NAVIGATING THE RAPIDS

Currently, there is no private market in basic flood insurance as the National Flood Insurance Program (NFIP) has preempted the field. Even if path dependence did not all but dictate continued government intervention, the development of a large market in unsubsidized, private flood insurance

¹¹² *Id.* (commenting that the bonus structure is not aligned with the NFIP goals of “increasing penetration in low-risk flood zones and among homeowners in all zones that do not have mortgages from federally regulated lenders”).

¹¹³ See Scales, *supra* note 6, at 33-34 (describing “disappointing” quality of help by adjusters in completing NFIP claims, which must be filed within sixty days of the loss).

¹¹⁴ See Gene Taylor, *Federal Insurance Reform after Katrina*, 77 MISS. L.J. 783, 786-87 (2008) (describing conflict and explaining that exacerbating the problem, at the instigation of the WYO companies, the NFIP implemented an expedited claims procedure after Katrina which allowed WYO companies to issue flood insurance checks “without apportioning the amount of wind and flood damage to structures with losses from both perils”). It also, however, became more difficult to obtain windstorm coverage in the aftermath of Katrina. *Id.* at 789-90. (Congressman Taylor did introduce legislation that would expand the NFIP to include windstorm.) See also KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 41-43 (describing the “wind-water controversy” and the Katrina-related lawsuits); Scales, *supra* note 6, at 24-29 (describing Katrina cases, including insurance companies’ interpretation of contract provisions yielding non-coverage for losses partially caused by flood and partially by wind).

¹¹⁵ See *id.* at 787-88 nn.14-15 (discussing press accounts in the Biloxi *Sun Herald* and *Times Picayune*); see also Scales, *supra* note 6, at 36-37 (describing an insurer’s “unusually attractive opportunity to recharacterize wind losses as flood losses as it is the very entity tasked with investigating flood claims for the government.”).

is in doubt.¹¹⁶ A private insurer would have to navigate multiple obstacles in setting a price that would be both actuarially sound and profitable.¹¹⁷ That price would almost certainly be viewed as too expensive by many individuals,¹¹⁸ including those who would be most in need of assistance following a flood.¹¹⁹ The first section of this Part reviews those pricing obstacles, including the extent to which universal coverage could alleviate those pressures. In addition, the section discusses the concern that universal coverage could increase moral hazard problems, including negative environmental externalities. The second section of this Part focuses on the consumer side of flood insurance and explores the puzzling reality that, even at subsidized rates, many individuals fail to plan for flood loss by purchasing insurance.

A. PROVIDER PERILS

Three well-known obstacles complicate the provision of flood insurance: correlation, adverse selection, and moral hazard. Universal

¹¹⁶ See MOSS, *supra* note 19, at 262 (describing failed private flood insurance experiments of the 1890s and 1920s); Jerry & Roberts, *supra* note 10, at 857 (arguing that “major disasters . . . require significant federal involvement for response and recovery”).

¹¹⁷ See Boardman, *supra* note 19, at 828 (“The primary problem for flood insurance is cost, not calculation.”); Scales, *supra* note 6, at 7 (explaining that flood insurance “suffers from unusual demand- and supply-side constraints that make it a relatively difficult market for insurers, and they have responded rationally by avoiding it”).

¹¹⁸ See *infra* Part III.B for a discussion of possible explanations rooted in cognitive psychology; see also Johnson et al., *supra* note 11, at 239 (explaining that flood loss risks are “underestimated systematically by homeowners in hazard-prone areas” and that residents will perceive “actuarially ‘fair’ coverage” as “overpriced, and will remain uninsured”).

¹¹⁹ See Debra Lyn Bassett, *Place, Disasters, and Disability*, in LAW AND RECOVERY FROM DISASTER: HURRICANE KATRINA 51, 64-69 (Robin Paul Malloy ed., 2009) (discussing rural poverty, including the “vulnerability of the rural disabled”); Klein & Zellmer, *supra* note 4, at 1473 (“Too often, those who suffer most are the poorest members of society.”); Levmore & Logue, *supra* note 4, at 317 (“Inner-city property owners, including businesses and homeowners, self-insure far more than their counterparts in affluent areas, in part because of availability problems.”); Kenneth B. Nunn, *Still Up on the Roof: Race, Victimology, and the Response to Hurricane Katrina*, in HURRICANE KATRINA: AMERICA’S UNNATURAL DISASTER 183, 184-87 (Jeremy I. Levitt & Matthew C. Whitaker, eds., 2009).

coverage should provide some relief as to the first and essentially sidestep the second. Moral hazard is more complicated, and expanded coverage would likely trigger concern that such coverage increases moral hazard problems, including environmental impacts.

1. Correlation

Flood losses are typically highly correlated.¹²⁰ That is, they generally occur simultaneously for a large swath of individuals. Thus, even if it is scientifically well established that a particular area suffers from a 1 in 100 chance of a flood in any particular year,¹²¹ if this year happens to be the year, all of the losses will occur at once. An insurance company may not yet have established sufficient reserves through receipt of premiums to cover the losses.¹²² Insurance companies operating within more limited geographic areas could face an even more concentrated correlation problem.¹²³

In order to deal with a correlation problem, a commercial insurance company would have to charge front-loaded premiums to create a large reserve in case the low probability event occurred early in the life of the

¹²⁰ See Jerry & Roberts, *supra* note 10, at 843 (explaining that flood risks are “difficult risks” because they are resistant to diversification and are highly correlated).

¹²¹ Complete statistical accuracy is, in fact, unlikely given the state of current flood maps. See *supra* Part II.C. Such ambiguity would likely further increase the premium. See Howard Kunreuther et al., *Insurer Ambiguity and Market Failure*, 7 J. RISK & UNCERTAINTY 71, 72 (2003) (describing survey data revealing that ambiguity in either probability of a loss or amount of loss results in “recommended premiums” that are “considerably higher”); Scales, *supra* note 6, at 8 (discussing ambiguity premium).

¹²² For-profit insurers will create insurance pools only if the contingencies are statistically predictable with respect to the pool as a whole but occur randomly with respect to any one contributor. The larger the pool of insureds, the more likely it is that the actuarial predictions will be sound and provide an adequate basis for calculating the premiums needed to cover the promised payouts and also yield a profit to the insurance company. See Jerry & Roberts, *supra* note 10, at 842-43 (describing insurance pools).

¹²³ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 65; Scales, *supra* note 6, 11 & n.30 (while cross-subsidization is possible, insurance companies oppose cross-subsidies whether between geographically distinct subsidiaries or between types of insurance (e.g., auto subsidizing casualty)); see Scales, *supra* note 6, at 11 (even national insurance companies generally operate through separate subsidiary companies organized along state lines or even smaller geographic regions).

risk pool. For example, if a commercial insurance company sought to create a pool for a flood plain subject to a 1 in 500 chance of a flood in any particular year, the premiums to establish the reserve would have to be high even during the early years of the contract in case the current year happened to be the year in which such a flood occurred.¹²⁴ Not only would individuals be unlikely to want to buy insurance requiring high up-front payments, they would also have such a low probability of receiving any payout during their lifetime that they would have a difficult time perceiving any benefit from the coverage.¹²⁵ Self-insurance would be the general choice.¹²⁶

Federal, universal coverage does not, of course, change the pattern of flood loss. It does, however, allow for greater diversification across geographic regions and access to non-program resources in particularly turbulent years. Even with a national program, flood losses can overtake capacity. This is essentially what happened to the NFIP during the 2005 hurricane season.¹²⁷ The NFIP met its obligations through its access to other resources — namely, its borrowing authority.¹²⁸

2. Adverse Selection

In addition to the need to price for correlation, an insurance company issuing a hypothetical flood loss policy would also have to price for a significant adverse selection problem. Adverse selection occurs when too many of the individuals who purchase coverage do so with certain or

¹²⁴ With thanks to David Cay Johnston for this example. *See Scales, supra* note 6, at 11 (explaining that correlation “induces greater variability in losses, leading to significantly higher premiums” if an insurance company is even willing to underwrite such a risk).

¹²⁵ *See infra* Part III.B, for a fuller discussion of consumer choices regarding flood preparation; Jerry & Roberts, *supra* note 10, at 845 (explaining that “having no claim” is often viewed as “purchasing a product with little value, notwithstanding that the person received security against loss”).

¹²⁶ Proposals to subsidize self-insurance have also been made. For example, Congress has proposed the creation of catastrophe savings devices — similar to health savings devices. *See Christine L. Agnew, Come Hell and High Water: Can the Tax Code Solve the Post-Katrina Insurance Crisis?*, 11 LEWIS & CLARK L. REV. 701, 738-43 (2007), for a critique of such an approach.

¹²⁷ *See supra* Part II.C.

¹²⁸ *See supra* note 79 and accompanying text.

near-certain knowledge that they will be filing an insurance claim.¹²⁹ For example, individuals will be more likely to purchase flood insurance if they have knowledge that the risk of flood loss is already at the doorstep (or roof, as the case may be).¹³⁰ Generally, the problem of adverse selection is one of information asymmetry.¹³¹ With respect to health and life insurance, this information asymmetry is fairly easy to conceptualize: the insurance company will not be privy to the private aches and pains of the insured and may under-price premiums as a result.¹³² In the case of floods, individuals would have particularized knowledge about the likelihood of flooding at a residence, and such knowledge would contribute to a classic adverse selection problem.

Adverse selection is a common reason advanced for the failure of a private flood insurance market to develop.¹³³ Universal or mandatory coverage is the classic solution to adverse selection.¹³⁴ If everyone is in the insurance pool, it removes the question of whether some are in the pool because they have inside information about personal risk. The information on flood risk developed through the NFIP, however, complicates the adverse selection picture. As discussed in Part II, part of the NFIP's mission is to assess flood risk and make those assessments available to the public. Thus, individuals can go to a FEMA website to look at flood risk maps.¹³⁵ Many of these maps are, as discussed in Part II, incomplete, difficult to decipher, or out of date, but, presumably, some will be influenced to purchase flood insurance as a result. Further, the lender

¹²⁹ See, e.g., Boardman, *supra* note 19, at 822 (“Adverse selection typically occurs when insurers cannot distinguish between higher and lower risk policyholders . . .”); see also Kaplow, *supra* note 10, at 543.

¹³⁰ A vivid example of such delayed response occurred during a flood in Chesterfield, Missouri, in 1993, when business property owners purchased flood insurance in response to a flood crest moving down the Missouri River. At the time, only a five-day waiting period was in place. See Klein & Zellmer, *supra* note 4, at 1493 (describing the event). Currently, a thirty-day waiting period applies. 42 U.S.C. § 4013(c) (2004); 44 C.F.R. § 61.11(c) (2010).

¹³¹ See Kaplow, *supra* note 10, at 543.

¹³² *Id.* at 545.

¹³³ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 135 (noting that private insurers argued that adverse selection required creation of the NFIP).

¹³⁴ See MOSS, *supra* note 19, at 50 (explaining that the ability of government to compel “broad participation” is “[p]erhaps the most widely recognized justification for public risk management”); Baker, *supra* note 20, at 380.

¹³⁵ Nat’l Flood Ins. Program, FLOODSMART.GOV, THE OFFICIAL SITE OF THE NFIP, <http://www.floodsmart.gov/floodsmart/> (last visited Aug. 25, 2011, 4:17 PM).

mandate applies only to high-risk property. Thus, the proportion of flood-prone properties among all the properties covered by the NFIP is likely high.¹³⁶ This result is not, however, readily ascribed to a classic adverse selection problem given that general flood risk information is primarily controlled and distributed by the government-insurer and is then used to enforce the lender mandate.¹³⁷

Private insurers would also have access to information about general flood risk and would, presumably, act in their own self-interest with two possible scenarios emerging. The first scenario assumes that demand is strongest among those with high-risk property and that as a result the insurance companies would have to charge higher premiums so as to account for high-risk property. Higher premiums could drive out lower-risk properties, necessitating premium increases, driving more lower-risk properties out — i.e., the replication of an adverse selection death spiral.¹³⁸ This cycle could prevent formation of a robust, private flood insurance option.¹³⁹ A second, arguably more plausible, possibility is that insurance companies would use their superior ability to assess risk to limit coverage only to those at lower risk of suffering damage in what has become known as a reverse information asymmetry problem.¹⁴⁰ As a result, higher-risk property would not be covered at all — a situation that would be incompatible with a goal of providing a stable flood loss safety net,

¹³⁶ See U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 14, at 5.

¹³⁷ See Michael Faure & Veronique Bruggerman, *Catastrophic Risks and First-party Insurance*, 15 CONN. INS. L.J. 1, 27 (2008) (under adverse selection information asymmetry “insurers must be unable to identify high-risk buyers”).

¹³⁸ See Kaplow, *supra* note 10, at 544; Scales, *supra* note 6, at 9 (suggesting that adverse selection “death spirals” occurring in the flood area is a possibility with “unique plausibility”). Cf. Faure & Bruggerman, *supra* note 137, at 26-27 (classic adverse selection “is not a serious problem” with respect to catastrophic losses); Peter Siegelman, *Adverse Selection in Insurance Markets: An Exaggerated Threat*, 113 YALE L.J. 1223 (2004) (refuting the long-held notion that adverse selection within insurance markets will inevitably lead to a collapse).

¹³⁹ Baker, *supra* note 20, at 378 (pointing out that this cycle illustrates that both insurer-side and insured-side adverse selection are at work).

¹⁴⁰ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 135 (describing that in the hurricane context, insurance companies may have the informational advantage “if insurance companies spend a lot of resources estimating the risk (which they do today)” and explaining that “[r]esearch . . . reveals that insurers might want to exploit this reverse information asymmetry, which results in low-risk individuals being optimally covered, while high-risk individuals are not”); see Baker, *supra* note 20, at 378.

though one that could lead to post-flood government intervention, at least as to dramatic flood events.¹⁴¹

3. Moral Hazard

Moral hazard is the term used for the notion that individuals will engage in cost-increasing behavior if they are able to shift some of the cost away from themselves.¹⁴² Since moral hazard is a potential side-effect of cost-shifting, moral hazard is a possible consequence of any opportunity for cost-shifting — whether insurance, post-disaster assistance, or even casualty loss tax deductions. While universal coverage helps solve the adverse selection problem, concerns about moral hazard could loom larger because of the increased opportunities for cost-shifting that would come with universal coverage.

An important assumption underlying the moral hazard concept is that an individual has a consistent cost tolerance with respect to a particular risk. If part of the cost has been shifted to another party, the benefitted individual will rationally engage in less careful behavior up until the point that the expected, unshifted costs reach that individual's tolerance threshold.¹⁴³ For example, a person with auto insurance would drive incrementally more recklessly than someone without insurance and, in theory, would set the level of additional recklessness so that any resulting damage would be adequately compensated by the policy and would not result in unanticipated, irreparable damage to person or property.¹⁴⁴

Insurers use various mechanisms to limit moral hazard, but the two most common monetary methods are co-pays and deductibles.¹⁴⁵ These

¹⁴¹ See *infra* Part III.B.

¹⁴² KENNETH BLACK, JR., & HAROLD D. SKIPPER, JR., *LIFE & HEALTH INSURANCE* 11 (13th ed. 2000).

¹⁴³ See Baker, *supra* note 21, at 270.

¹⁴⁴ See *id.* at 276-78 (explaining that an assumption underlying the economics of moral hazard is that “money compensates for loss” when in fact “money cannot restore the sense of security lost when a storm destroys a home . . . or, indeed, much of what is important in life”).

¹⁴⁵ See Boardman, *supra* note 19, at 841 (noting that “moral hazard is always tempered by the extent to which the policyholder remains on the risk, through deductibles, caps, and the uncertainty of a compliant insurer”); Johnson et al., *supra* note 11, at 232 (“The most common mechanism for controlling moral hazard is a deductible”); KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 99 (discussing NFIP deductibles and stating that “the majority of homeowners prefer a lower deductible”). The NFIP does use deductibles, but since the rates are not

devices are intended to shift just enough pain back to the individuals so that they are more reluctant to engage in the cost-increasing behavior. Even though co-pays and deductibles are usually quite small relative to the costs that are covered by the insurance policy, out-of-pocket costs are fixed, certain losses that individuals may be particularly prone to shun.¹⁴⁶ Indeed, setting co-pays too high may increase rather than decrease moral hazard by over-detering individuals from seeking benefits. For example, if an individual puts off medical care to avoid a co-payment, the cost of the later treatment may be much higher.¹⁴⁷

In addition to using the pain of out-of-pocket costs to control for moral hazard loss, insurers may also monitor the behavior of insured individuals and thereby require a particular level of care.¹⁴⁸ Direct observation of the day-to-day behavior of individuals can be costly, but for many types of coverage, insurance companies have devised methods for indirect monitoring, including reliance on monitoring devices (e.g., fire alarms) or third parties (e.g., doctors).¹⁴⁹ The NFIP requires community adherence to floodplain regulations to increase care and lower the costs of flood loss.¹⁵⁰ Premium rebates or adjustments could be used as monetary rewards for easily measured good behavior — e.g., an absence of claims on the policy.¹⁵¹

In the case of flood loss compensation for individuals, the primary moral hazard concerns arise with respect to how individuals store their personal possessions, how individuals construct and maintain their homes,

actuarially sound, these deductibles may not have the desired effect. The NFIP also limits payouts to the value of the damaged property instead of allowing for payment tied to replacement cost, unless the damage is to a primary residence and its contents. *See also supra* Part II.C.

¹⁴⁶ *See* RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 33-34 (2008) (discussing loss aversion). *See also infra* Part III.B (discussing how consumer cognitive perceptions affect flood loss coverage).

¹⁴⁷ ABHIJIT V. BANERJEE & ESTHER DUFLO, POOR ECONOMICS: A RADICAL RETHINKING OF THE WAY TO FIGHT GLOBAL POVERTY (2011).

¹⁴⁸ Baker, *supra* note 21, at 280-81.

¹⁴⁹ *Id.*

¹⁵⁰ *See supra* Part II.A.

¹⁵¹ *See* Johnson et al., *supra* note 11, at 232-33, 238; Baker, *supra* note 21, at 270 (discussing that for some types of moral hazard, observational monitoring is more critical — for example, if the insurance reduces “the incentive to minimize the cost of recovering from a loss,” e.g., the “malingering aspect of the disability insurance temptation problem”).

and where individuals choose to live. Coverage expansion would trigger concerns about exponentially increased moral hazard costs, particularly environmental costs associated with increased development. Expansion of social safety net coverage for individual homeowners and renters could, however, have less of an effect on moral hazard costs than may appear upon first consideration because the assumptions underlying moral hazard analysis are less likely to hold true as to social safety net coverage for primary residences.¹⁵²

In his work excavating the historical and theoretical landscape of moral hazard, Professor Tom Baker outlined several assumptions behind classic moral hazard analysis.¹⁵³ The realities of flood loss suggest that several of these assumptions do not hold true, particularly as to an individual's primary residence. Moral hazard analysis assumes that "money compensates for loss."¹⁵⁴ While loss of a vacation home may come close to being compensable by money, the loss of a primary home and its contents is far less likely to satisfy this condition.¹⁵⁵

Another assumption underlying moral hazard is that "people with insurance have control over themselves and their property."¹⁵⁶ Of course, individuals have some choice over where to live, but, for many individuals, such choices will be constrained by many factors, including financial and social. Further, in the case of flood loss, any particular individual is likely to be far removed from decisions involving flood plain regulation and

¹⁵² See Baker, *supra* note 21, at 240 ("By 'proving' that helping people has harmful consequences, the economics of moral hazard justify the abandonment of legal rules and social policies that try to help the less fortunate."); Kunreuther & Pauly, *supra* note 70, at 108 ("If consumers generally ignore both loss probabilities and potential government assistance in deciding whether or not to buy insurance and how much insurance to purchase, . . . [p]ublic intervention based on our concern for fellow citizens can be straightforward: provide as much assistance as our conscience dictates to fill in the observed gaps in coverage . . . If such choices represent outcomes that are incomplete or inefficient according to the 'selfish' expected utility model, it is irrelevant because people are not using this model of choice anyway."). *But see* Trebilcock & Daniels, *supra* note 10, at 104 (describing the "perverse incentive effects" of post-disaster relief as "severely exacerbating problems of adverse selection and moral hazard in locational decisions").

¹⁵³ Baker, *supra* note 21, at 276.

¹⁵⁴ *Id.*

¹⁵⁵ *See id.* at 276-78 ("[Money cannot restore the sense of security lost when a storm destroys a home . . . or, indeed, much of what is important in life."].

¹⁵⁶ *Id.* at 276.

development.¹⁵⁷ Expansion of flood insurance to all primary residences would potentially affect the care taken by residential developers and landlords,¹⁵⁸ but such effects could be handled directly rather than being used as a reason for denying social benefits to more vulnerable individuals.¹⁵⁹

Moral hazard analysis depends also on individuals being “rational loss minimizers.”¹⁶⁰ As will be discussed in greater detail in the next section, there is reason to believe that a great many individuals fail to act rationally with respect to flood loss. If individuals have difficulty understanding and planning for flood risk, they may also have trouble engaging in the calculated, care reducing behavior assumed by moral hazard analysis. Of course, some individuals will strategically engage in less careful behavior. For example, under the NFIP, the extent of repetitive loss, particularly for second homes,¹⁶¹ as well as the concentration of coverage in high-risk areas could suggest a moral hazard problem.¹⁶² But the concentration of policies in high-risk areas could also be attributable in part to the lender mandate¹⁶³ or to adverse selection.¹⁶⁴

The moral hazard effects of flood insurance expansion also depend on the extent to which post-disaster relief already stands in for universal coverage.¹⁶⁵ Post-disaster relief operates to shift risk and thus raises moral

¹⁵⁷ See *id.* at 279 (“If the people exposed to the insurance incentive are not in control of the behavior that matters, then reducing the insurance incentive will impose a cost on those people while providing little benefit . . .”).

¹⁵⁸ The problem of business flood loss coverage will be addressed in a subsequent article.

¹⁵⁹ See Baker, *supra* note 21, at 240 (“[C]onventional economic accounts of moral hazard exaggerate the incentive effects of real-world insurance and, at the same time, underestimate the social benefits of insurance.”).

¹⁶⁰ *Id.* at 276.

¹⁶¹ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 85 (The CBO “found that many subsidized properties in coastal areas (23 percent from their sample of 10,000 properties) were second homes, vacation homes, or rentals.”).

¹⁶² KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 93-94 (A study undertaken by Professors Kunreuther and Michel-Kerjan of the Florida market revealed that five counties in Florida accounted for two-thirds of the flood policies in Florida; these counties were coastal counties whereas the five counties with the lowest number of policies were located well inland.).

¹⁶³ See *supra* Part II.B.

¹⁶⁴ See *supra* Part III.A.2.

¹⁶⁵ See Pasterick, *supra* note 31, at 152 (“The prevailing public impression is that federal disaster assistance is generally equivalent to the financial protection provided by hazard insurance. In reality this is not the case.”).

hazard concerns similar to those of *ex ante* coverage.¹⁶⁶ Post-disaster relief for large flood events is virtually guaranteed,¹⁶⁷ and even for smaller scale events, various tax provisions operate to shift some of the risk.¹⁶⁸ As with flood insurance coverage, the moral hazard story for post-disaster assistance also depends, however, on assumptions that may not hold true for flood loss. For example, the patchwork nature¹⁶⁹ of available post-disaster relief may make being a “rational loss minimizer” even more difficult.¹⁷⁰

Given the history of flood loss in the United States, there seems little doubt that more care should be taken in land use and development.¹⁷¹ At the same time, however, it is less clear the extent to which classic moral hazard analysis satisfactorily explains the problem, particularly if the focus is on individual homeowners and renters. Even if flood loss protection does not fit neatly into a classic moral hazard frame, the problem of unwise, environmentally harmful development remains. The inability of individuals to plan carefully for flood loss suggests that steps for greater care, including not only mitigation but prohibitions, must be express and be backed by strong incentives or even mandates. Expansion of social safety net coverage could provide an opportunity to craft such incentives and to

¹⁶⁶ See Levmore & Logue, *supra* note 4, at 281 (“[E]xpectation of federal relief has almost certainly increased the willingness of some individuals and businesses to locate or remain in disaster prone areas.”).

¹⁶⁷ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 122 (“[T]he driving force in the provision of government assistance, is the occurrence of large-scale losses.”).

¹⁶⁸ See *infra* Part V.

¹⁶⁹ See *infra* Part III.B (discussion of problems associated with post-disaster relief).

¹⁷⁰ Baker, *supra* note 21, at 276; see also KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 122 (Empirical work on post-disaster relief suggests that “individuals or communities have not based their protective decisions in advance of a disaster by focusing on the expectation of government assistance.” Professors Kunreuther and Michel-Kerjan cite studies suggesting that “most homeowners in earthquake- and hurricane-prone areas did not expect to receive aid from the federal government following a disaster” and that “local governments that received disaster relief undertook more efforts to reduce losses from future disasters than those who did not.” Professors Kunreuther and Michel-Kerjan conclude “this behavior seems counterintuitive, and the reasons for it are not fully understood.”).

¹⁷¹ See generally Klein & Zellmer, *supra* note 4; KLEIN & ZELLMER, *supra* note 7.

enlist homeowners and renters in reducing harm caused by developers and other real property businesses, such as landlords.¹⁷²

B. DEMAND AND ITS DISCONTENTS

The central demand puzzle is why so many homeowners and renters fail to purchase or under-purchase flood insurance, even though it is a bargain. Examples of this puzzle can be gleaned from news accounts of recent flooding. In June 2011, the Souris River rose and caused massive flooding in Minot, North Dakota.¹⁷³ The river had previously seemed nonthreatening after numerous public works initiatives had reduced flood risk.¹⁷⁴ In 2000, the federal government had moved the flood risk assessment level outside the high risk category, which meant that lenders no longer had to enforce the mandate to purchase flood insurance.¹⁷⁵ Although residents remained eligible to participate in flood insurance and were counseled by federal officials to maintain their policies, a large number dropped coverage.¹⁷⁶ At the time of the flooding, an estimated one in ten had flood insurance.¹⁷⁷ In 2011, only 476 residents had flood insurance policies; just one year earlier, 959 residents had flood insurance.¹⁷⁸ The combination of public works projects, lowered risk assessment, removal of the mandate, and financial pressures inexorably led individuals to stop worrying about floods.¹⁷⁹ As one resident put it, “I didn’t have any concerns. . . . It was not going to happen to me. I was in complete denial.”¹⁸⁰

¹⁷² See MOSS, *supra* note 19, at 50-51 (“[G]overnment enjoys a considerable advantage over private insurers when it comes to monitoring and controlling moral hazard directly.”).

¹⁷³ A.G. Sulzberger, *They Dropped Their Flood Insurance, Then the ‘Mouse’ Roared*, N.Y. TIMES, June 24, 2011, at A13.

¹⁷⁴ *Id.* (“[T]he once flood-prone river—known locally as the Mouse, after its French name—had seemingly been tamed by public works projects that reshaped the channel, raised the banks and controlled the flow of water . . .”).

¹⁷⁵ *Id.*; see *supra* Part II.B (discussing lender mandate).

¹⁷⁶ See Sulzberger, *supra* note 171.

¹⁷⁷ *Id.*

¹⁷⁸ *Id.*

¹⁷⁹ *Id.* (“[A]nother problem facing residents of Minot is a consequence not of failing to control the river but of decades of doing so successfully. . . . ‘Some citizens have been lulled into a false sense of security because we have had such good results,’ said . . . the City Council president.”).

¹⁸⁰ *Id.* (statement by a real estate agent married to a firefighter).

This response to the possibility of flood loss is not unusual. Even though the NFIP provides flood insurance at low rates,¹⁸¹ many individuals still do not purchase it.¹⁸² The study of financial preparedness, including the problem of underinsurance,¹⁸³ has increasingly become intertwined with cognitive considerations such as optimism bias, loss aversion, and time-inconsistent preferences.¹⁸⁴ This section briefly reviews some of the potential contributions of this research to the under-purchase of flood insurance.

¹⁸¹ See Johnson et al., *supra* note 11, at 225 (“A rational, risk-neutral consumer would purchase coverage at an actuarially fair price that is equivalent to the expected loss. . . . In practice, the story is apparently not that simple.”).

¹⁸² See Johnson et al., *supra* note 11, at 225 (noting that “coverage is underpurchased by consumers, even when it is heavily subsidized”); Kunreuther & Pauly, *supra* note 70, at 103 (“The NFIP . . . provides highly subsidized rates for existing homes so that any risk-averse individual who made the appropriate calculations of the expected benefits and costs of purchasing such insurance should have wanted coverage. In the Louisiana parishes affected by Katrina the percentage of homeowners with flood insurance ranged from 57.7 percent . . . to 7.3 percent. . . .”).

¹⁸³ Underinsurance is a problem for virtually all potentially financially devastating events — for example, death, disability, and casualty. See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 16 (noting that thirty-eight percent of “owner-occupied homes with severe wind damage” in the 2005 hurricanes did not have insurance against wind loss); Levmore & Logue, *supra* note 4, at 273-74 (discussing problem of underinsurance for life insurance after the attacks of 9/11); Francine J. Lipman, *Anatomy of a Disaster Under the Internal Revenue Code*, 6 FLA. TAX REV. 953, 972-73 (2005) (describing fire underinsurance in California).

¹⁸⁴ See, e.g., THALER & SUNSTEIN, *supra* note 146, at 101-56 (discussing cognitive glitches and financial decisions); Levmore & Logue, *supra* note 4, at 282-83 (stating that “simple underinsurance” may result from “myopia, overoptimism, bad planning, or passivity.”); Tom C.W. Lin, *A Behavioral Framework for Securities Risk*, 34 SEATTLE U. L. REV. 325, 336-40 (2011) (discussing the rational investor versus the real investor); Edward J. McCaffery & Joel Slemrod, *Toward an Agenda for Behavioral Public Finance*, in BEHAVIORAL PUBLIC FINANCE 3,13 (Edward J. McCaffery & Joel Slemrod, eds., 2006) (discussing application of “time-inconsistency models” to savings decisions); Robert J. Meyer, *Why We Under-Prepare for Hazards*, in ON RISK AND DISASTER: LESSONS FROM HURRICANE KATRINA 153, 154-68 (Ronald J. Daniels et al., eds., 2006) (discussing inference bias, forecast bias, procrastination, status quo bias, and empathy gaps); Scales, *supra* note 6, at 9-10 (explaining individuals’ tendencies to respond differently to risks that they view as remote).

Individuals appear to have difficulty conceptualizing probabilities.¹⁸⁵ For low probability events that carry large costs, individuals often fail to take minimal, economically rational steps — purchasing flood insurance, for example. On the other hand, many individuals over-pay for insurance for events that have more salience — e.g., warranties for small electronics¹⁸⁶ or flight insurance following acts of or warnings about terrorism.¹⁸⁷ Using familiarity as a shortcut for understanding a given probability may work relatively well in a variety of situations¹⁸⁸ but is problematic for flood events.¹⁸⁹ Even individuals residing in a relatively hazardous area may never have personally experienced a flood event.¹⁹⁰

¹⁸⁵ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 121 (discussing studies suggesting that people cannot “distinguish between probabilities that ranged from 1 in 10,000 to 1 in 1 million” and that individuals also “did not respond to insurance premiums as a signal of risk”); Jerry & Roberts, *supra* note 10, at 845 (discussing lack of demand for coverage of difficult risks as relating to whether the individual has “past experience with it or know someone else who has endured it”); Johnson et al., *supra* note 11, at 225-26 (explaining that “consumers may have distorted perceptions of the size or probability of the risks they face.”). *See also supra* Part III.A.3 (discussing assumption of accurate risk assessment underlying moral hazard analysis).

¹⁸⁶ THALER & SUNSTEIN, *supra* note 144, at 78-80 (discussing extended warranties on small devices and concluding “the extended warranty is a product that simply should not exist” given various market assumptions).

¹⁸⁷ *See* Johnson et al., *supra* note 11, at 226-31 (discussing “distorted beliefs concerning the probability and size of some potential losses” following from vivid and dramatic news events, including terrorism). *See also* KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 122 (discussing study finding that “local governments that received disaster relief undertook more efforts to reduce losses from future disasters than those who did not”).

¹⁸⁸ THALER & SUNSTEIN, *supra* note 146, at 24-26 (discussing cluster of related mental shortcuts tied to familiarity, including the availability heuristic, accessibility and salience).

¹⁸⁹ *See* Kunreuther & Pauly, *supra* note 70, at 106-07 (discussing how “[r]ather than using the expected utility model, many residents in hazard prone areas appear to follow a sequential model of choice” and “[f]or these individuals only after the occurrence of a disaster does this event assume sufficient salience”). For example, the purchase of NFIP policies increased dramatically following the 2005 hurricane season. KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 87 (750,000 more policies at end of 2007 than in 2005).

¹⁹⁰ *See* Kunreuther & Pauly, *supra* note 70, at 105 (characterizing a “hazard-prone area” as one where annual probability of damage “is within the range of 1 in 50 to 1 in 500. So, while the financial losses should such an event occur can be

Individuals may also be overly optimistic when faced with probabilistic information.¹⁹¹ Thus, even assuming individuals spend the time needed to understand flood risk,¹⁹² such information may still not be enough to overcome an optimistic feeling that the event will not actually happen. As discussed in Part II and also alluded to in the anecdote beginning this section, public works projects may further contribute to a false sense of security.¹⁹³ Individuals who initially purchase a policy may later cancel because of difficulty in perceiving the benefits of a policy that has not produced a cash transfer to the insured.¹⁹⁴ Flood insurance coverage may seem superfluous to an individual who has paid for the coverage for many years but who has yet to file a claim.¹⁹⁵ Individuals already feeling budget constraints will be more prone to seeing the coverage as a luxury rather than necessity.¹⁹⁶ Structuring insurance covering low probability

significant, the great majority of people will not have observed an event close at hand recently.”).

¹⁹¹ See, e.g., THALER & SUNSTEIN, *supra* note 146, at 32-33 (discussing “[u]nrealistic optimism” with respect to statistical risks “to life and health”); Lin, *supra* note 184, at 340 (“Despite facts to the contrary, individuals generally have an overabundance of confidence in their own abilities and an overabundance of optimism in their futures.”).

¹⁹² As discussed *supra* even expert agencies have difficulty creating and maintaining accurate flood risk maps. See Kunreuther & Pauly, *supra* note 70, at 105 (“[M]any potential victims of disaster perceive the costs of getting information about the hazard and costs of protection to be so high relative to the expected benefits that they do not even consider purchasing insurance.”) (citation omitted).

¹⁹³ See generally *supra* Part II.A (discussing the unintended consequences public works projects may have). See also Sulzberger, *supra* note 173 (“Some residents said they had misinterpreted these revised flood estimates to mean that they were no longer at risk. Others said they had just used the lower odds as an opportunity to save some money.”).

¹⁹⁴ See Johnson et al., *supra* note 11, at 231-35 (discussing framing effects and the relative attractiveness of rebates over deductibles). See also THALER & SUNSTEIN, *supra* note 146, at 36-37 (discussing framing effects and “choice architects”).

¹⁹⁵ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 124 (“People often purchase flood insurance only after suffering damage in a flood, but many cancel their policies when several consecutive years pass with no flood.”); Kunreuther & Pauly, *supra* note 70, at 107 (stating that there is “empirical evidence that many homeowners who initially purchase insurance are likely to cancel policies if they have not made a claim over the course of the next few years”); Scales, *supra* note 6, at 31 n.108 (“[U]nrealized insurance risks still have substantial value.”).

¹⁹⁶ Kunreuther & Pauly, *supra* note 70, at 105-06 (“[R]eluctance to invest in protection voluntarily is compounded by budget constraints. For some

events so that it pays an annual rebate to individuals who have not filed a claim may help increase policy retention.¹⁹⁷

Even if individuals understand that buying insurance would be economically wise,¹⁹⁸ they may decide to wait until tomorrow to make the purchase given the pain of parting with money today.¹⁹⁹ Unfortunately, individuals tend to keep moving that “tomorrow” forward in time until it becomes too late.²⁰⁰ Possible contributors to the procrastination phenomenon include an aversion to parting with cash in exchange for uncertain benefits²⁰¹ and a bias toward maintaining one’s current position.²⁰²

homeowners with relatively low incomes, disaster insurance is considered a discretionary expense. . . . In contrast to the expected utility model where the demand for insurance depends on the premium relative to the expected loss, demand appears to depend only on the premium for a given amount of coverage.”).

¹⁹⁷ Johnson et al., *supra* note 11, at 233-35 (describing experiment suggesting that disability insurance structured to provide rebates would be more attractive than standard disability insurance). *See also* BANERJEE & DUFLO, *supra* note 145, at 62-65 (describing how making transfers of small amounts of food supplies increased participation in vaccination program — a program that required multiple treatments and would yield protection benefits that would occur in the future and be difficult to perceive).

¹⁹⁸ The difficulty individuals have in understanding probabilities and coverage benefits will reinforce the desire to procrastinate. *See* BANERJEE & DUFLO, *supra* note 147, at 154 (“[T]he [procrastination] problem is made even harder when the insurance is against a catastrophic event: The payout would take place . . . in a particularly unpleasant future that no one really wants to think about.”); Meyer, *supra* note 184, at 164 (“Decisions to invest in protection against low-probability events are particularly susceptible to procrastination . . .”).

¹⁹⁹ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 122 (explaining that “some homeowners with relatively low incomes” will perceive disaster insurance as a “discretionary expense that should be incurred only if residual funds are available after taking care of what individuals or families consider to be the necessities of life”).

²⁰⁰ BANERJEE & DUFLO, *supra* note 147, at 65 (“Our natural inclination is to postpone small costs, so that they are borne not by our today self but by our tomorrow self instead.”); *see also* Richard H. Thaler & Shlomo Benartzi, *Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving*, 112 J. POL. ECON. S164, S167-68 (2004) (discussing the concepts of self-control and procrastination).

²⁰¹ *See, e.g.*, THALER & SUNSTEIN, *supra* note 146, at 33-34 (describing loss aversion); Thaler & Benartzi, *supra* note 200, at S169-70 (describing loss aversion on savings behavior).

²⁰² THALER & SUNSTEIN, *supra* note 146, at 34-35 (discussing status quo bias).

Pre-commitment devices may solve some types of procrastination problems.²⁰³ The tax system already yields examples of such devices. Congress codified²⁰⁴ an administrative position²⁰⁵ through which employers may enroll employees in section 401(k) deferred compensation plan by default; employees who do not want to participate must then complete an opt-out procedure. Although employees may fairly easily free themselves from their bindings, inertia will likely keep most from doing so and will thereby reduce future regrets over poor planning.²⁰⁶ In addition to this example of a congressionally crafted technique, numerous individuals save through the tax system by selecting or sticking with tax withholding rates

²⁰³ See McCaffery & Slemrod, *supra* note 184, at 13 (describing advantages of “self-commitment devices that limit future choices, like Ulysses did when he bound himself to the mast as his ship passed the Sirens’ sweet song.”). The extent to which government should act paternalistically to remedy cognitive flaws is strongly debated. See, e.g., THALER & SUNSTEIN, *supra* note 146, at 4-11; see generally Richard A. Epstein, *Second-Order Rationality*, in BEHAVIORAL PUBLIC FINANCE 355 (Edward J. McCaffery & Joel Slemrod, eds., 2006); Jonathan Klick & Gregory Mitchell, *Government Regulation of Irrationality: Moral and Cognitive Hazards*, 90 MINN. L. REV. 1620 (2006); Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 NW. U. L. REV. 1165 (2003); Cass R. Sunstein & Richard H. Thaler, *Libertarian Paternalism Is Not an Oxymoron*, 70 U. CHI. L. REV. 1159 (2003).

²⁰⁴ Pension Protection Act of 2006, Pub. L. No. 109-280 § 902, 120 Stat. 780, 1033. See also Automatic Contribution Arrangements, 74 Fed. Reg. 8200, 8200-02 (Feb. 24, 2009); Notice 2009-65, 2009-39 I.R.B. 413 (Sept. 28, 2009) (sample plan amendments for adding § 401(k)(13) automatic enrollment features); see generally STAFF OF THE J. COMM. ON TAXATION, GENERAL EXPLANATION OF TAX LEGISLATION ENACTED IN THE 109TH 329 (Jan. 17, 2007).

²⁰⁵ Rev. Rul. 2000-8, 2000-1 C.B. 617, *amplifying & superseding* Rev. Rul. 98-30, 1998-1 C.B. 1273.

²⁰⁶ See THALER & SUNSTEIN, *supra* note 146, at 107-09; James J. Choi et al., *Saving for Retirement on the Path of Least Resistance*, in BEHAVIORAL PUBLIC FINANCE 304, 339 (Edward J. McCaffery & Joel Slemrod, eds., 2006) (discussing evidence that employees make savings decisions passively and arguing that “employers should choose their plan defaults carefully, since these defaults will strongly influence the retirement preparation of their employees”). See also STAFF OF THE J. COMM. ON TAXATION, PRESENT LAW AND ANALYSIS RELATING TO INDIVIDUAL RETIREMENT ARRANGEMENTS 51-52 (June 26, 2008). (“The theory is that to the extent that these employees are not saving for retirement due to inertia (simple failure to take initiative), that same failure to take initiative may prevent them from electing out of the contributions” and will thereby assist “employees who can and want to save for retirement.”).

that yield significant, lump sum refunds.²⁰⁷ Even though saving through withholding seems to make little economic sense because of the foregone interest, the technique helps individuals resist the temptation to spend the money elsewhere while providing an easy, virtually painless path to amassing a usefully large sum.²⁰⁸

In the case of disaster insurance, devices for dealing with lack of preparation may need to be stronger given the difficulties associated with processing flood loss probabilities.²⁰⁹ The costs of failure to take mitigation steps may make a disaster more costly,²¹⁰ yet the more costly the more likely it is that aftermath aid will be provided. As discussed in the previous section, adverse selection also presents a problem against which mandates provide significant protection. The adverse selection problem could be

²⁰⁷ Fennell, *supra* note 26, at 148 (“About three-fourths of U.S. taxpayers have more income tax than necessary withheld . . . or make excess estimated payments. . .”) (internal citation omitted).

²⁰⁸ STUART RUTHERFORD, *THE POOR AND THEIR MONEY* 1-7 (2009) (discussing need and ways poor amass “usefully large lump sums”). *See also* Fennell, *supra* note 26, at 148-52 (exploring explanations for over-withholding preference, including its use as a pre-commitment device). The allure of lump sums may inspire other techniques designed to combat under-saving. Recently, for example, some U.S. credit unions are attempting to correct savings myopia by adding a lottery hook. *See* Melissa Schettini Kearney et al., *Making Savers Winners: An Overview of Prize-linked Savings Products* 14-20 (Nat’l Bureau of Econ. Research, Working Paper No. 16433) (2010), *available at* www.nber.org/papers/w16433 (discussing U.S. market potential and current offerings); Anne Stuhldreher, *Credit Unions Launch a Savings Lottery, and Everyone Hits the Jackpot*, WASH. POST, Feb. 7, 2010, at B4 (discussing savings lotteries). Such lottery-linked accounts have been utilized internationally for years. *See* Mauro F. Guillén & Adrian E. Tschoegl, *Banking on Gambling: Banks and Lottery-Linked Deposit Accounts*, 21 J. FIN. SERVICES RES. 219, 225-29 (overview of history, practice, and methods used in various countries); *See generally* Kearney et al., *supra* note 208, at 7-14 (discussing use of programs used internationally).

²⁰⁹ Kunreuther & Pauly, *supra* note 70, at 103 (discussing evidence suggesting that people’s beliefs about flood loss cause them to “have no incentive to invest in protective measures voluntarily”).

²¹⁰ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 262-63 (discussing “natural disaster syndrome” as increased vulnerability caused by “cost-effective loss-reduction measures” and reviewing “extensive evidence that residents in hazard-prone areas do not undertake loss prevention measures voluntarily”).

exacerbated by cognitive hurdles if flood loss is salient only for those most at risk.²¹¹

Local, state, and federal officials attempting to plan for over-optimism and probability processing difficulty will themselves be subject to the same types of cognitive challenges.²¹² Prior to a flood, government actors may fail to take protective steps even though cost-benefit analysis strongly supports action.²¹³ Political pressures to limit spending and keep taxes low may further dampen efforts to take precautionary measures.²¹⁴ Yet, in the aftermath of a flood, especially a large-scale event, officials will

²¹¹ It is also possible, however, that the problem might be lessened if even individuals facing the highest risk fail to take action because of various cognitive hurdles. Further, if individuals only perceive flood loss as salient after an event occurs, adverse selection may be lower because another event in the near future may be less likely depending on community response. See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 122 (discussing study finding “that local governments that received disaster relief undertook more efforts to reduce losses from future disasters than those who did not”).

²¹² See Meyer, *supra* note 184, at 173 (“[B]enevolent central planning” is limited in “that it has legitimacy only to the degree that benevolent central planning is free of the decision biases that it is meant to cure.”); Scales, *supra* note 6, at 12 (“Governments, like individuals, are subject to many of the cognitive biases that constrain the development of private catastrophe insurance.”).

²¹³ Burby, *supra* note 38, at 179 (providing three examples of how local government (in)action in New Orleans revealed a lack of concern about flooding hazards, including lobbying by the local government for levees built to resist a one-hundred-year flood rather than a two-hundred-year flood in order to reduce the local cost share); Kunreuther & Pauly, *supra* note 70, at 102 (“Public sector agencies may also behave in ways that are inconsistent with optimal social policy by not using the principles of benefit-cost analysis . . . as illustrated by the Corps of Engineers decision not to strengthen the New Orleans levees.”); Meyer, *supra* note 184, at 157 (discussing history of hurricanes in the greater New Orleans area and noting that “ironically, this success [with Hurricane Camille]—combined with the lack of storms in the years that followed—seemed to deflate rather than spur interest in completing the [flood-control] project.”); Nunn, *supra* note 119, at 186-90 (detailing information available to public officials regarding the vulnerability of New Orleans).

²¹⁴ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 263 (discussing how “given short-term reelection considerations, the representative is likely to vote for measures that allocate taxpayers’ money elsewhere that yield more political capital. . . . because they believe that their constituents are not worried about these events occurring”).

be required to do something²¹⁵ and may reap political rewards for their public acts of generosity.²¹⁶ Reliance on *ex post* relief may carry with it significant problems. Relief efforts will depend on the vividness of the event—and, with respect to government assistance, may also depend on the proximity of the event to an election.²¹⁷ If the event is sufficiently large scale, aid may be relatively plentiful.²¹⁸ On the other hand, even if a flood event is catastrophic in the life of a particular family, if the flood is an isolated occurrence, that family may have little access to outside sources of support.²¹⁹ Even in cases of large-scale disasters where aftermath aid is relatively plentiful, access to the aid may be difficult for individuals to obtain because the path may not be clear having been put together in a patchy, *ad hoc* fashion in a stressful context.²²⁰ Lower-income individuals may suffer in particular. For example, an important post-disaster program is the availability of low-interest loans from the Small Business Administration for damaged property, including personal residences and

²¹⁵ *Id.* at 262 (“The magnitude of the destruction following a catastrophe often leads public sector agencies to provide disaster relief to victims even if the government claimed it had no intention of doing so prior to the event.”). *See also*, Daniel Shavero, *Beyond Public Choice and Public Interest: A Study of the Legislative Process as Illustrated by Tax Legislation in the 1980s*, 139 U. PA. L. REV. 1, 86-87 (1990) (discussing congressional “bias in favor of action over inaction”).

²¹⁶ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 263 (“The fact that politicians can benefit from their generous actions following a disaster raises basic questions as to the capacity of elected representatives at the local, state, and federal levels to induce people to adopt protection measures before the next disaster.”).

²¹⁷ *See* KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 123 (describing research showing that “disaster assistance is more prevalent in presidential election years, all other things being equal”); Kunreuther & Pauly, *supra* note 70, at 106 (“[T]he amount and terms of the disaster [relief] depend on random political influences including the proximity of the disaster to the date of the next national election.”).

²¹⁸ Kunreuther & Pauly, *supra* note 70, at 106 (“What is well understood is that large-scale losses from disasters are a driving force with respect to the actual provision of government relief (citation omitted) . . .”).

²¹⁹ *See* Ellen P. Aprill & Richard Schmalbeck, *Post-Disaster Tax Legislation: A Series of Unfortunate Events*, 56 DUKE L.J. 51, 61-62 (2006) (discussing horizontal inequity in comparing relief for large-scale disasters and that for disasters affecting fewer individuals).

²²⁰ *See* Kunreuther & Pauly, *supra* note 70, at 106 (“[T]he combination of low private insurance and haphazard public disaster relief may lead to inefficiency as well as high levels of government spending.”).

effects.²²¹ Low-income individuals are often ineligible for these loans because of the default risk.²²²

Costs may be higher with post-disaster assistance—in part because the cost of administering and obtaining the aid may be more costly because of lack of pre-planning and in part because the costs may be higher than if adequate pre-disaster mitigation steps had occurred.²²³ In the aftermath of a disaster, the government may overreact by enacting rules that are inconsistent with other policy goals — tax changes, for example, that have far larger effects than may have been intended.²²⁴ Of course, *ex ante* provisions are unlikely to bring the need for aftermath aid down to zero. Unanticipated problems may emerge and some coverage gaps may remain.

²²¹ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 19 (describing program).

²²² *Id.* at 19. See also BANERJEE & DUFLO, *supra* note 145, at 151-52 (discussing government intervention in international context and noting “[t]he government intervenes only in cases of large-scale disasters, not when a buffalo dies or someone is hit by a car. And even disaster relief is, in most cases, vastly insufficient by the time it gets to the poor.”).

²²³ KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 262 (noting that the “combination of underinvestment in protection prior to the event leading to large disaster losses, together with the general taxpayer financing some of the recovery, can be critiqued on both efficiency and equity grounds”).

²²⁴ See Danshera Cords, *Charitable Contributions for Disaster Relief: Rationalizing Tax Consequences and Victim Benefits*, 57 CATH. U. L. REV. 427, 434 (2008) (concluding that “Congress should avoid post-disaster temporary tax legislation as a means to aid disaster relief efforts”); Aprill & Schmalbeck, *supra* note 219, at 53-54 (discussing Congressional overreaction and the “legislative imperative” to act following a disaster and concluding that the results have “been disappointing, and largely inconsistent with sound tax policy”).

Professors Ellen Aprill and Richard Schmalbeck, for example, have recommended having Congress adopt joint resolutions declaring a disaster instead of delegating to the executive branch the responsibility of designating federally declared disasters because “Congress will likely always feel that it needs to act when disaster strikes.” Aprill & Schmalbeck, *supra* note 219, at 95. They have also recommended creation of a panel to identify categories of relief provisions—some of which would be available widely and other that should rarely be used. *Id.* at 97-99. Such “[g]uidelines . . . would establish presumptions, obligating a member of Congress who proposes to disregard them to offer compelling explanations of why it would be appropriate to do so.” *Id.*

In addition, government officials, charities, and individuals will likely still want to do something to show altruism and support.²²⁵

No simple solution exists to deal with the difficulties inherent in flood loss and floodplain management. The approach proposed in this Article is one that relies on having multiple pressure points for action and needed adjustment with respect to flood loss.

IV. NATIONAL FLOOD LOSS SECURITY PROGRAM

The previous two parts outlined some of the reasons supporting the case for continued government intervention in flood loss relief and for structuring such intervention to be widely available and focused on limiting *ad hoc*, post-disaster decisions. Much more could (and has) been written on these issues. This section will, however, take as a working assumption that the benefits of a broad, *ex ante* approach outweigh its costs and will turn to discussing the potential benefits of structuring a national flood loss security program using the powerful tools available through the tax system. This Part also outlines one possible structure for such an approach.²²⁶ Part V

²²⁵ Cf. Levmore & Logue, *supra* note 4, at 277 (predicting that “public and charitable relief will more likely be forthcoming if there is (or is perceived to be) less than full private insurance”).

²²⁶ The mechanisms proposed in this Article are aimed directly at individuals instead of being designed to have an effect on institutions potentially involved in managing flood risk — e.g., insurance companies and charitable organizations providing aftermath aid. Thus, for example, this Article does not include discussion of possible subsidies for insurance companies to aid in the creation of a commercial flood insurance market. See Agnew, *supra* note 124 (discussing proposed legislation aimed at providing tax relief to insurance companies for catastrophe reserves). Nor does it include discussion of some type of “supercharged subsidy for charitable gifts”. See Levmore & Logue, *supra* note 4, at 308-09 (discussing such a proposal in the context of terrorism insurance).

The money to fund flood loss coverage could also be raised through a consumption tax model. State sales taxes are examples of a consumption tax; excise taxes on alcohol and cigarettes are federal examples of consumption taxes. See JOEL SLEMRD AND JON BAKIJA, *TAXING OURSELVES: A CITIZEN’S GUIDE TO THE DEBATE OVER TAXES* 231-68 (4th ed. 2008) (discussing consumption taxes). The rate of a consumption tax would, however, be much more difficult, if not impossible, to tie to a particular individual’s flood risk. It would also be more difficult to adjust consumption tax rates to take into account an individual’s ability to pay. For example, imagine that a flood tax were imposed as a national sales tax; to adjust for flood risk and ability to pay, at each point of sale, a questionnaire regarding one’s income and location of principal residence would need to be

discusses how current tax law on disaster relief should be adjusted so as to harmonize with the creation of a broad flood loss security program.

A. LEVERS OF TAX SYSTEM POWER

Flood loss protection is highly complex and requires attention to both social safety net concerns and concerns regarding unsafe or unwise development and construction. Utilizing tax system components to implement flood loss protection could provide multiple avenues for addressing this complexity. Use of the tax system would facilitate implementation of mandates and universal coverage, thus ensuring a minimum level of coverage for all citizens. Universal coverage would also help to alleviate adverse selection problems and to resolve the difficulty individuals have in committing to flood loss prevention. Other tax system components — refunds and rate adjustments, for example — could be utilized to make the benefits of having coverage more salient and to incentivize individuals to engage in mitigation efforts. The tax system could also be structured so as to harmonize with and reinforce other flood-cost reduction programs, including relocation programs.

The strength of the withholding mechanism would facilitate the collection of premiums.²²⁷ Other tax return items — gross income, for example — could be readily utilized to adjust premiums so as to take into account an individual's ability-to-pay. As was discussed in Part II, premium collection is currently outsourced to private insurance businesses with highly problematic results. The IRS, in contrast, has a strong record of enforcement competence and general efficiency.²²⁸ Further, the IRS and

completed. While the process could be streamlined through technology — e.g., a smart card — the administrative and compliance problems of using a sales tax for such a purpose loom large.

A more realistic consumption tax approach would utilize a low-rate consumption tax to support a supplemental general catastrophe fund for dealing with unexpected costs. Such a fund could also provide a focal point for political involvement in the aftermath of the disaster. See Aprill & Schmalbeck, *supra* note 219, at 93.

²²⁷ SLEMROD & BAKIJA, *supra* note 226, at 181 (discussing how withholding technique is a “major enforcement tool”). But see Richard L. Doernberg, *The Case Against Withholding*, 61 TEX. L. REV. 595 (1982) (discussing history of withholding system and providing a critique of the system).

²²⁸ John T. Scholz, *Contractual Compliance and The Federal Income Tax System*, 13 WASH. U. J.L. & POL'Y 139, 162 (2003) (“Efficiency pressures are so embedded in the organizational culture of the IRS that even strong external

Treasury already have experience dealing with flood events as it must enforce several tax rules relating to natural disasters.²²⁹

Of course, bringing in the IRS and Treasury will also raise new concerns. Utilizing these governmental units to implement a social program could further dilute their mission, particularly revenue collection under the income tax system.²³⁰ The IRS and Treasury already play a significant role in other social programs, such as retirement planning and health care. In addition, the Internal Revenue Code contains numerous tax expenditures and other indirect social programs, such as the earned income tax credit. The detrimental effects of the addition of one more social program to be administered in part by the IRS and Treasury is hard to know in advance. Certainly, implementation of the proposed flood loss security program would require expansion of the IRS budget, something that is politically difficult even in less partisan times. On the other hand, if the goal is universal coverage through a federal program, it is difficult to envision a government agency or private organization better equipped to handle the collection of premiums.

The IRS and Treasury would not be the only administrative agencies tasked with overseeing the proposed program. Flood risk assessment and oversight of community regulations would still belong to the agency that currently handles those assessments — i.e., FEMA.²³¹ In addition, FEMA's role would need to expand to include claims adjustment, a function which is currently almost entirely outsourced to private insurance companies through the WYO program.²³² The proposed program's heavy reliance on administrative agencies raises concerns

pressures have limited ability to change them.”); John T. Scholz & B. Dan Wood, *Efficiency, Equity, and Politics: Democratic Controls Over the Tax Collector*, 43 AMER. J. POL. SCI. 1166, 1184-85 (1999) (finding that “efficiency consistently provides the dominant influence on audit allocation decisions”). Complaints about the IRS being *too* driven by collection may, however, arise as they have in the past. Scholz, *supra* at 164-65 (discussing efforts by Congress to discourage “unduly zealous enforcement”).

²²⁹ See *infra* Part V.

²³⁰ For general discussion of IRS mission and history, see Alan H. Plumley & C. Eugene Steuerle, *Ultimate Objectives for the IRS: Balancing Revenue and Service*, in THE CRISIS IN TAX ADMINISTRATION 311 (Henry J. Aaron & Joel Slemrod, eds., 2004).

²³¹ See *supra* Part II.C (discussing problems with current FEMA maps).

²³² See *supra* Part II.D.

regarding agency capture and other agency shortcomings.²³³ As discussed in Part II, a case can already be made based on the history of the NFIP that communities exert too much influence over the updating and enforcement of new flood maps, and FEMA's approach to Hurricane Katrina can be cited as a textbook example of regulatory failure.²³⁴ While an in-depth discussion of agency capture and other potential agency flaws is beyond the scope of this Article, the proposed system arguably should not be any more problematic than that under the current NFIP and may even be less susceptible to such pressures.

Moving to a mandatory, universally applicable system may make interest group formation more difficult.²³⁵ Under the NFIP, communities opt in to the program, and the availability of flood insurance to individuals depends on communities agreeing to participate in the NFIP. FEMA appears to have as an internal goal a focus on individual access and purchase of flood insurance.²³⁶ If that is the case, FEMA may be more inclined to agree to community demands in order to facilitate that mission since individual access is available only if the community qualifies as an NFIP participant.²³⁷ If flood insurance purchase is mandatory for individuals, access to coverage would not be held hostage by community demands. The rates charged to individuals under the system proposed in this Article would, however, be adjusted through community adherence to regulations. Thus, pressure from communities on agencies would continue to be a factor, but the issue of rate rather than access may be less likely to

²³³ See MUELLER, *supra* note 40, at 343-47 (discussing phenomenon of rent-seeking through regulation).

²³⁴ See *supra* Parts II.A and C. See also Russell S. Sobel & Pater T Leeson, *Government's Response to Hurricane Katrina: A Public Choice Analysis*, 127 PUBLIC CHOICE 55 (2006).

²³⁵ See MUELLER, *supra* note 40, at 475 ("One of the most counterintuitive predictions of Olson's theory is that small interest groups are much more effective at obtaining favors from government than large groups are. . . . In poor countries, where the agricultural sector is large and the group of middle-class urban dwellers is small, farmers receive small or even *negative* subsidies for their products . . . [but if] farmers make up a tiny fraction of the total workforce, they often receive giant subsidies.").

²³⁶ See *The Official Site of the National Flood Insurance Program*, FLOODSMART.GOV, www.floodsmart.gov.

²³⁷ See Robert M. Howard & David C. Nixon, *Local Control of the Bureaucracy: Federal Appeals Courts, Ideology, and the Internal Revenue Service*, 13 WASH. U. J.L. & POL'Y 233, 236 (2003) (discussing the "goal-seeking nature of bureaucracy").

induce capitulation to community pressures. (Of course, communities may themselves be under greater pressure from their residents as a mandate may mean that more individuals would take an interest in assuring community compliance so as to receive the best premium rates possible.)

The involvement of the IRS and Treasury may act as a counterweight to community pressure and provide monitoring of FEMA.²³⁸ Some scholarship suggests that the IRS and Treasury are less susceptible to capture than other agencies because of the diverse range of interests in the charge of these agencies.²³⁹ In addition, empirical research on IRS enforcement patterns suggests that the IRS is more influenced by national trends than by localized politics.²⁴⁰

²³⁸ See Rachel E. Barkow, *Insulating Agencies: Avoiding Capture Through Institutional Design*, 89 TEX. L. REV. 15, 49-58 (2010) (exploring how “shared responsibilities” among agencies “can either foster or frustrate” agency independence).

²³⁹ Scholz, *supra* note 228, at 158-59 (“Of all of the specialized enforcement agencies, the IRS is arguably the most sheltered from direct political influence at all levels.”); Edward A. Zelinsky, *James Madison and Public Choice at Gucci Gulch: A Procedural Defense of Tax Expenditures and Tax Institutions*, 102 YALE L.J. 1165, 1166-67 (1993) (“Tax institutions, because of their greater visibility and more competitive nature, are less susceptible to interest group capture and possess greater legitimacy under pluralist criteria than their direct expenditure equivalents.”).

²⁴⁰ Howard & Nixon, *supra* note 237, at 233 (“Examining cross-sectional time series data from 1960 until 1988, we found that the IRS shifts the number of audits it conducts of businesses versus individuals in response to the prevailing median ideology of the federal courts of appeals, and in response to the prevailing ideological framework of the President and Congress.”); Scholz & Wood, *supra* note 228, at 1185 (“Partisan responsiveness exerts a somewhat less consistent influence on audit allocations. State-level partisanship consistently shifts audit resources away from taxpayers with business income in Republican states, but the results are less supportive of the partisanship hypothesis for nonbusiness taxpayers. On the national level, both presidents and Congressional committees influence the tradeoff between equity and efficiency, with presidential influence being significant for more categories of taxpayers than committee influence.”); John T. Scholz & Dan Wood, *Controlling the IRS: Principals, Principles, and Public Administration*, 42 AM. J. POL. SCI. 141, 160 (1998) (“Consistent with past research on other agencies, the mix of IRS audits also responds to changes in the presidency as well as changes in the leadership and ideology of members of congressional oversight committees. On the other hand, the mix of corporate versus individual audits does not respond to state-level variations in partisanship of the state’s congressional delegation, governor, presidential vote, or legislature . . .

Given the recent series of congressional showdowns over deficits, social programs, and taxes, enactment of such an expansion over the current NFIP would face its own hurdles. During the last several years, Congress has put off dealing with the shortcomings of the NFIP by enacting short-term extensions of the program.²⁴¹ Admittedly, the prospects of a more complete overhaul of the program are relatively dim given the current political climate. Of particular concern may be the mandatory aspect of the proposed expansion,²⁴² especially given the litigation surrounding the mandate contained in the health care legislation.²⁴³ Discussion of the constitutionality of the health care mandate is beyond the scope of this Article, but there is reason to think that the structure proposed in this Article is less susceptible to such arguments.

The federal government already has a well-established commercial interest in flood loss protection as evidenced by the NFIP, Army Corps of Engineers flood mitigation projects, and the provision of aftermath protection.²⁴⁴ The formation and presence of commercial special interest groups should also be much lower than was the case with health insurance given that private insurers have not underwritten flood insurance for decades, although removal of the WYO payments may cause some consternation.²⁴⁵ The collection of premiums would be somewhat similar to that utilized for social security, a program whose constitutionality has been

The picture suggests that earlier reforms have succeeded in insulating field offices from local influences.”).

²⁴¹ See *supra* Part I.

²⁴² See Kunreuther & Pauly, *supra* note 70, at 114 (discussing the prospects of flood insurance mandate and suggesting “Lower income people will have the increases cushioned (though not taken away entirely) by subsidies, but the middle class especially may object to being charged for insurance which they think they do not need and will never use. How to assemble at least a minimal winning coalition of citizens to make mandated coverage feasible is a crucial research topic.”).

²⁴³ See *Thomas More Law Center v. Obama*, 720 F. Supp. 2d 882 (E.D. Mich. 2010), *aff’d* 108 A.F.T.R.2d 2011-5007 (6th Cir. 2011); *Florida ex rel Bondi vs. Health & Human Services*, 107 A.F.T.R.2d 2011-724 (N.D. Fla. 2011), *aff’d in part, rev’d in part*, 2011 WL 3519178 (11th Cir. 2011).

²⁴⁴ See 42 U.S.C. § 4001-4002 (2006) (congressional findings regarding economic burdens caused by flood loss). For discussion of the commerce clause as it relates to the healthcare legislation, compare *Thomas More Law Center* 108 A.F.T.R. 2d 2011-5007 at 14-15, with *Florida ex rel Bondi*, 2011 WL 3519178 at *24-38.

²⁴⁵ See *supra* Part II.D (discussion of WYO program).

upheld.²⁴⁶ The premiums would be paid in substantial part as an exchange for direct coverage rather than being a penalty related to a decision to self-insure (which has been characterized by critics of the health care mandate as a tax on doing nothing rather than an income or excise tax).²⁴⁷ The vividness of recent flood events and the feelings of altruism triggered by such events may also ease the path to enactment.²⁴⁸ Finally, it may be possible to invest a portion of the collected revenues (in years of lower flooding costs) to spur development of private catastrophe coverage — for example, stimulation of the catastrophe bond market.²⁴⁹

²⁴⁶ See *Steward Machine Co. v. Davis*, 301 U.S. 548 (1937) (employer portion). For an overview of the current state of the social security system, see Patricia Dilley, *Through the Doughnut Hole: Reimagining the Social Security Contribution & Benefit Base Limit*, 62 ADMIN. LAW REV. 367 (2010).

Although the proposed plan uses the term “premium,” the payments could also be characterized as a form of income tax under an analysis applied to social security as well as to the “shared responsibility payment” of the health care legislation. See Brian Galle, *Conditional Taxation and the Constitutionality of Health Care Reform*, 120 YALE L.J. ONLINE 27 (2010) (responsibility payment is a constitutional income tax); Edward Kleinbard, *Constitutional Kreplach*, TAX NOTES 755, 761-62 (Aug. 16, 2010) (the healthcare penalty is a constitutional income tax and one tied to self-insurance). But see Steven J. Willis & Nakku Chung, *Constitutional Decapitation and Healthcare*, TAX NOTES 169 (July 12, 2010) (arguing that penalty is an unconstitutional, unapportioned direct tax — assuming it is a tax).

Individuals may feel less favorably towards taxes and penalties and more favorably toward rewards, even if the two structures are economically identical. Individuals also appear to prefer hidden taxes to obvious taxes. See George Lowenstein et al., *Statistical, Identifiable, and Iconic Victims*, in BEHAVIORAL PUBLIC FINANCE 32, 38-39 (Edward J. McCaffery & Joel Slemrod, eds. 2006) (discussing the appeal of hidden taxes). In the case of the health care legislation, using a term that avoided the word “tax” was viewed as disingenuous and backfired. Use of the term premium should be less problematic in the case of flood loss protection given that it is paid in exchange for coverage.

²⁴⁷ See Willis & Chung, *supra* note 246, at 185 (“Congress could require everyone to purchase flood insurance from the government and charge appropriately for it.”). See also Kleinbard, *supra* note 246, at 759 (explaining that “[T]he Supreme Court has rejected any invitation to distinguish between taxes designed to influence behavior and taxes designed to raise revenue.”).

²⁴⁸ See Kunreuther & Pauly, *supra* note 70, at 108 (“Concern for our fellow citizens as well as our own needs should disaster strike home makes us want our government to help out, and in a democracy the public sector responds.”).

²⁴⁹ With thanks to Yariv Brauner & Tom Lin for this suggestion. For discussion of catastrophe bonds and other alternative risk transfer instruments, see

While the costs and hurdles to enactment of universal flood loss will remain largely unknown until such a program is put into place, the costs to individuals and communities of continuing with the NFIP and the *ad hoc* post-disaster relief are relatively well understood. While this Article advocates a universal system, if such a system were politically impossible a scaled-back version of the system proposed herein could still be a significant improvement over the current approach to flood loss.

B. PROGRAM OUTLINE

Payment into the proposed flood loss security program would be mandatory for individuals,²⁵⁰ and premium collection would be handled as much as possible through withholding, with adjustments as necessary through an individual's annual income tax return. Calculating the withholding rate could be simplified by making various default assumptions, which could be then be adjusted through worksheets completed with the annual income tax return.²⁵¹ Preferably, the default withholding rates should be set so that is more likely that individual adjustments lead to a refund rather than to the requirement of additional

KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 174-90; *see also* Scales, *supra* note 6, at 46.

²⁵⁰ As discussed *supra* Part IV.B, for individuals at low risk, the tax could be made an opt-out program if universal coverage were too politically difficult to enact. *Supra* note 24. Using an opt-out regime rather than opt-in would allow for the strategic use of the status quo bias, as has been allowed for 401(k) plans. *See supra* Part III.B. Such opting out could come at the price of losing certain other tax benefits, such as the casualty loss deduction. *See infra* Part V.C.

²⁵¹ Complicated details relating to filing status—e.g., married filing jointly—would have to be worked out, and that level of detail is beyond the scope of this project. Working out those details may, however, be smoothed by similarities to other withholding programs. For example, the flood security tax system would share similarities with the current system for withholding regular income taxes and the requirement for estimated payments. *See* Doernberg, *supra* note 227, at 595 (discussing history of withholding system and providing a critique of the system). Self-employed individuals are also required to remit self-employment tax with their tax form each year. *See* Patricia Dilley, *Breaking the Glass Slipper: Reflections on the Self-Employment Tax*, 54 TAX LAW. 65, n.6 (2000). For a discussion of the conceptual flaws surrounding the self-employment tax, *see* Patricia Dilley, *Breaking the Glass Slipper: Reflections on the Self-Employment Tax*, 54 TAX LAW. 65 (2000).

payments.²⁵² The premium rate would depend on the flood risk loss, the amount of coverage purchased, and ability to pay.

The flood risk assessment would be tied to the location of the principal residence.²⁵³ Second homes would not be covered, which should help curb repetitive loss problems and is also in keeping with an approach focused on provision of a safety net.²⁵⁴ Thus, it will be critical to define principal residence carefully.²⁵⁵ The tax code already uses this term in other contexts,²⁵⁶ and the same basic approach as contained in those sections could be utilized. Thus, an individual's principal residence would depend on various factors, including place of employment, length of abode, and residence of family members. Ownership would not be required,²⁵⁷ though coverage would then, of course, be limited to possessions. Some individuals may have difficulty pointing to a principal residence — either

²⁵² See *supra* Part III.B (discussion of individual preferences for tax refunds).

²⁵³ A rate that varies with location raises the question whether the Uniformity Clause would present an obstacle to enactment of such a program. The Uniformity Clause is contained in Article I, section 8, which provides “The Congress shall have Power To lay and collect Taxes, Duties, Imposts and Excises . . . but all Duties, Imposts and Excises shall be uniform throughout the United States.” See Aprill & Schmalbeck, *supra* note 219, at 78-84 (discussing the uniformity clause); Lawrence Zelenak, *Are Rifle Shot Transition Rules and Other Ad Hoc Tax Legislation Constitutional?*, 44 TAX L. REV. 563, 588-601 (1989). This Article's proposals should pass muster under the *Ptasynski* case. *United States v. Ptasyński*, 462 U.S. 74 (1983). In that case, Congress imposed an excise tax on crude oil that varied according to three tiers and that also exempted “Alaskan oil,” which was defined in terms of a well's proximity to the Arctic Circle or Alaska-Aleutian Range and Trans-Alaska Pipeline. *Id.* at 77-78. The Court explained in dictum, “[h]ad Congress described this class of oil in nongeographic terms, there would be no question as to the Act's constitutionality.” *Id.* at 86. See Zelenak, *supra* at 591-94 (explaining significance of this dictum and arguing that Supreme Court is likely to apply it in future cases). The Court upheld the exemption even though it was framed in geographic terms because “Congress has exercised its considered judgment with respect to an enormously complex problem.” *Ptasynski*, 462 U.S. at 86.

²⁵⁴ As will be discussed in greater detail *infra*, coverage could also be designed so as to limit repetitive claims with respect to the same structure. See *infra* notes 277-79 and accompanying text.

²⁵⁵ A procedure for changing the primary residence would have to be put in place as well.

²⁵⁶ I.R.C. §§ 121, 123, 1033(h) (2006).

²⁵⁷ The regulations promulgated under Code section 123 have a similar provision. Treas. Reg. § 1.123-1(c) (as amended in 1980).

because they have two or more regular residences or because they have no residence at all. Regulations issued under an unrelated provision provide that taxpayers with more than one residence are generally treated as having their primary residence as the place where they spend the most time.²⁵⁸ In the case of flood coverage, in limited circumstances,²⁵⁹ it may make sense to allow taxpayers to designate a principal residence.²⁶⁰

Once the principal residence has been identified, the flood risk associated with the principal residence would have to be determined. This determination clearly presents an administrative burden, but it is one that is already present even if the current system takes a less visible approach through the WYO program²⁶¹ and the lender mandate.²⁶² Risk rate brackets would be created, and these brackets could be narrowly or loosely tailored. One possibility is to mimic the current approach under the NFIP and use broad designations. For example, three brackets — high risk, moderate risk, and low risk — could be used as an initial matter. The high-risk category would apply to homes in one- hundred year flood plains or greater risk, which corresponds to the current high-risk designation in the NFIP.²⁶³ The moderate risk category could apply to homes facing a five- hundred year flood plain risk or greater (but less than the one- hundred year flood risk).²⁶⁴ All other homes would be low risk.

As discussed in Part II, flood risk assessments have not been completed (or are badly in need of updating) for many communities. Individuals with principal residences in such areas would still need to be assigned to a risk category. Default assignment to the high-risk category could maximize the possibility that flood risk assessment would be completed since the individual would have an incentive to pursue completion of the assessment. It could also forestall complaints about being

²⁵⁸ Treas. Reg. § 1.121-1(b) (as amended in 2002).

²⁵⁹ For example, designation could be freely allowed for high-risk residences but subject to much greater scrutiny if the designation relates to a home in a lower-risk area.

²⁶⁰ It would be possible for each spouse in a marriage to have a separate principal residence if, for example, each spouse has a different home for purposes of the “away from home” requirement of section 162. *See* I.R.C. § 162 (2006). Care would be required to keep such an allowance from becoming a means to circumvent the principal residence requirement.

²⁶¹ *See supra* Part II.D.

²⁶² *See supra* Part I.B.

²⁶³ *See supra* note 59 (discussion of term).

²⁶⁴ *See supra* note 62 (describing recommendation for 500-year flood plain).

moved from moderate-risk to high-risk.²⁶⁵ At the same time, individuals may view an assignment to such a category would undoubtedly be viewed as punitive by many individuals; thus, it may be politically prudent to set the default for unmapped areas to moderate risk.

Use of risk rate brackets could function as an incentive for individuals to lower their risk rate by engaging in less risky behavior (or by influencing their communities to meet guidelines that would also move the flood plain risk). In theory, if individuals have a choice of moving to a high-risk or moderate-risk primary residence, all other things being equal, they should choose the moderate-risk home to lower the taxes. The brackets could be used in other ways to minimize costly behavior. For example, an individual who experiences a flood loss and receives a payment under the program could automatically be moved into a higher risk category until the individual shows proof of taking adequate mitigation²⁶⁶ or relocation to a less risky principal residence.

Rebates could be used to reward individuals who engage in hazard mitigation or have multiple years without a claim. As discussed in Part III, individuals appear to prefer to have taxes over-withheld so as to receive the lump-sum tax rebate payment,²⁶⁷ and individuals may also prefer insurance rebates (coupled with higher base insurance rates) to deductibles.²⁶⁸ Because flood loss is relatively unlikely even for individuals residing in high-risk zones,²⁶⁹ interim rewards through refunds may help ease the psychic difficulty of contributing to a system that in most years may be perceived as not providing a benefit.²⁷⁰ With a national, mandatory program, individuals may be more likely to understand the probability of flood loss because flood losses, if looked at using a national perspective, may appear more salient.²⁷¹

²⁶⁵ See *supra* Part II.C (discussing how FEMA has adopted a grandfathering approach in response to such complaints).

²⁶⁶ Of course, mitigation devices are themselves not without risk. See *supra* Part II.A. See also Klein & Zellmer, *supra* note 4, at 1486-89 (discussing the inadequacies of “engineered flood control”).

²⁶⁷ Fennell, *supra* note 26, at 148-52.

²⁶⁸ Johnson et al., *supra* note 11, at 232-33, 238.

²⁶⁹ See *supra* Part III.B.

²⁷⁰ See *supra* Part III.B.

²⁷¹ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 352 (“[W]hen one expands the lens to include a state or country or the global community, catastrophic risks have a much higher likelihood of occurring.”).

Community involvement in flood mitigation would remain a part of the proposal as mitigation does reduce flood losses (within limits).²⁷² Homes in nonparticipating communities could automatically be treated as being in a high-risk area, while communities that receive high mitigation ratings could trigger rate reductions for their residents. Thus, the Community Rating System, described in Part II, would remain an important feature of the flood loss landscape.

Risk would not be the only item to affecting rate, and adjustments would also be made for coverage and income. In order for the program to function as a social safety net, minimum coverage levels as well as maximum coverage levels would need to be set. The minimum coverage level should be tied to local cost of living measures. The maximum coverage limits under the current NFIP appear generally adequate.²⁷³ These limits are \$100,000 for personal property and \$250,000 for residential real estate.²⁷⁴ The coverage would apply per residence, so a married couple sharing the same principal residence would have the same coverage limits as a single individual residing alone in one principal residence. Above the minimum coverage level, individuals would be required to demonstrate actual loss rather than receiving the replacement value amount.²⁷⁵ A side effect of the proposed flood security plan may be a decrease in the aftermath relief provided by private sources and through special legislation.²⁷⁶ Thus, minimum coverage should include payments for temporary living expense grants.

²⁷² See *supra* Part II.A. See also Burby, *supra* note 38, at 182 (“The number of NFIP insurance claims per capita for compensation of flood damages and the per capita dollar amount of payments made to settle claims were *highest* in states that did not require responsible behavior—neither building code enforcement nor comprehensive plans—from their local governments. . .”).

²⁷³ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 83 (describing study of years 2000-2005 suggesting that “almost three-quarters were still below the \$250,000 maximum coverage limit. One reason for this large percentage is that many homes had property values below this limit.”).

²⁷⁴ See *supra* Part II.C. This coverage will not, of course, provide a full recovery for many residences. Individuals with residences worth in excess of the maximum coverage would be left to seek excess coverage in the private market, to the extent available. KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 41 (noting that in Katrina some homes covered by flood insurance still suffered large uninsured losses because of the \$250,000 NFIP cap and failure to obtain “excess coverage from private carriers”). Such a result is, however, consistent with the safety-net focus of the proposed program.

²⁷⁵ See *supra* note 75.

²⁷⁶ See Levmore & Logue, *supra* note 4.

One possibility for further tamping down repetitive loss would be to have structural coverage run with the property rather than with the individual²⁷⁷ and limit the recovery per property to a particular number of times, through a declining coverage regime, or through a combination of the two. For example, maximum structural coverage could be reduced in half to \$125,000 for a second occurrence, halved again for a third occurrence, with coverage disappearing entirely for a fourth occurrence.²⁷⁸ Such a system would add some further complication to the collection system, and notations would also need to be added to deeds so that purchasers would not be caught unawares. The threat of coverage removal would also have to be credible.²⁷⁹ Coupling coverage reductions with relocation grants may be advisable as may be providing some type of reset mechanism in the event of community changes.

The amount of tax owed would also be adjusted for income level — with “income” tied to gross income rather than to “wages”.²⁸⁰ Tax-exempt interest should be added back in for a more accurate snapshot of an individual’s ability to pay.²⁸¹

Because adjusting for income levels would further complicate the proposed withholding system, it may be advisable to have fairly broad categories and then create credits for the poorest individuals. For example, the withholding rate could remain unchanged from \$1 to \$250,000, from \$250,000 to \$999,999, and finally from \$1 million and up.²⁸² Lower income

²⁷⁷ With thanks to Marty McMahon for this suggestion. *See also* Scales, *supra* note 6, at 20 n. 70 (noting that “insurance does not ‘run with the land’” in discussing lender mandate since “mortgage obligations have a life of their own”).

²⁷⁸ *See* KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 264-65 (“In areas that have suffered multiple catastrophes—say, three or more—nature may be telling us something: that these locations are naturally much more likely to be damaged than others.”).

²⁷⁹ *See* Kydland & Prescott, *supra* note 45, at 477 (“But the rational agent knows that, if he and others build houses there, the government will take the necessary flood-control measures.”).

²⁸⁰ The definition of “wage” may be quite complex. I.R.C. § 3401 (2006).

²⁸¹ A similar rule applies to the taxation of Social Security benefits. I.R.C. § 86(b)(2)(B) (2006). *See* Goldin v. Baker, 809 F.2d 187, 188 (2d Cir. 1987) (upholding constitutionality of § 86(b)(2)(B)).

²⁸² By comparison, the rate brackets in the general income tax system are more compressed and the highest rate bracket begins at a fairly low level. *See* Martin J. McMahon, Jr., *The Matthew Effect and Federal Taxation*, 45 B.C. L. REV. 993 (2004) (discussing distribution of income tax system brackets).

individuals could then receive credits to further assist them in participating in the system.

Adjusting the premium for wealth rather than for gross income would arguably provide a more accurate picture of an individual's ability to pay the tax, particularly since the coverage would be for a wealth loss, but measuring wealth would be far more difficult than measuring income given that there is no annually assessed U.S. wealth tax.²⁸³ Coverage levels may, in any case, be a rough proxy for wealth. That is, wealthier individuals may be more likely to seek to cover the maximum amount of property damage, and the rate can be increased for larger coverage amounts. As will be discussed in the next Part, coverage limitations should be enforced directly but also indirectly through, for example, limitations on the casualty loss deduction.

V. CHANGES TO THE CODE

Various provisions of the Internal Revenue Code provide additional risk-shifting from individuals to the government (and then out to other citizens).²⁸⁴ This section outlines the current tax treatment of: non-insurance benefits received from government or private actors; insurance proceeds for property loss and for temporary assistance; and losses not reimbursed by insurance, government, or other private actors.²⁸⁵ As to each

²⁸³ For a discussion of wealth taxes, see Joseph M. Dodge, *What Federal Taxes Are Subject to the Rule of Apportionment under the Constitution?*, 11 U. PA. J. CONST. L. 839 (2009); David Shakow & Reed Shuldiner, *A Comprehensive Wealth Tax*, 53 TAX L. REV. 499 (2000).

²⁸⁴ See KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 19-20 (explaining that federal tax policy on catastrophe losses "affect the risk mitigation incentives of property owners and insurers' ability to finance catastrophe losses"); MOSS, *supra* note 19 (describing various ways governments intervene in regulating risk).

²⁸⁵ The discussion in the Article centers on those Code provisions aimed most directly at individuals and their personal property losses, but Congress has in the past enacted and may again enact other special relief rules in the event of a disaster, including provisions aimed at business losses. See generally, James Edward Maule, *Tax Incentives for Economically Distressed Areas*, in BNA TAX MANAGEMENT PORTFOLIO no.597 (2007). Congress may enact business-related provisions—e.g., enhanced expensing or net operating loss treatment. State and local governments' ability to issue bonds may be increased and restrictions on certain credits, such as the low income housing credit, may be lifted. See I.R.C. §§ 1400L-1400Q (2006). Further, charities and charitable deductions may receive favorable treatment. See I.R.C. §§ 1400L-1400Q. Penalties on retirement account withdrawals may be lifted and deadlines extended for various tax items. See I.R.C.

group, this section also discusses changes that may be recommended so as to harmonize these provisions with the proposed flood loss security program.²⁸⁶ Such harmonization is achieved through favorable tax treatment for benefits received under the proposed program and supporting mitigation grant programs while placing some limits the tax benefits to be obtained for non-program assistance and strongly limiting the deductibility of uncompensated flood losses.

A. NON-INSURANCE ASSISTANCE

In the immediate aftermath of a flood, government agencies, charitable organizations, commercial businesses and individuals frequently provide temporary aid to the victims. This aid is likely to include fresh water, meals, hygiene supplies, clothing, transportation, and shelter.²⁸⁷ From a traditional, economic approach to defining income, such items are arguably taxable increases to the recipients. Not surprisingly given the circumstances in which these transfers occur, the value of temporary aid for disaster victims is generally excluded from taxable income, though until about ten years ago, the path for exclusion depended in large part on Service rulings²⁸⁸ and was sometimes arguably inconsistent with the Internal Revenue Code.²⁸⁹

§§ 1400L-1400Q. *See also* Aprill & Schmalbeck, *supra* note 219 (providing overview of relief enacted in response to 2005 hurricane season and attacks of September 11, 2001); Lipman, *supra* note 183, at 976-1018 (describing Code sections aimed at disaster relief).

²⁸⁶ Even in the absence of enactment of the proposed expansion of flood insurance, these tax sections could be better aligned with the goals of the NFIP and other flood-related programs. *See* KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 20 (“[C]urrent tax policy with respect to uninsured disaster losses has received little attention to date, as it creates disincentives for efficient disaster risk management.”).

²⁸⁷ Disaster relief grants made to businesses are not addressed in this Article. For background on such grants, *see* Notice 2003-18, 2003-1 C.B. 699 (grants to businesses affected by World Trade Center attacks not excludable as gifts or as general welfare payments).

²⁸⁸ For example, under these older authorities, if temporary assistance came from government, it would be treated as a nontaxable, general welfare distribution. Rev. Rul. 98-19, 1998-1 C.B. 840; Rev. Rul. 76-144, 1976-1 C.B. 17,18. Not all governmental transfers are excluded from gross income. For example, unemployment is included because it is substitute for wages. I.R.C. § 85 (2006). *See also* Rev. Rul. 85-29 (Alaska dividend payments are income); J. MARTIN BURKE & MICHAEL K. FRIEL, *TAXATION OF INDIVIDUAL INCOME* (9th ed. 2010), at

225-26 (discussing general welfare rulings). In some cases, use of the general welfare exclusion was technically problematic. For example, in the aftermath of a fire caused by the National Park Service, the Service had difficulty determining whether relief payments that were also a settlement of any claims against the federal government could qualify under the general welfare exclusion and whether a distinction should be drawn between insured and uninsured individuals. *See infra* Part V.B.1 (discussing Code section 123 which provides a limited exclusion for payments under insurance contracts for temporary living expense assistance). The Chief Counsel's office recommended not taxing any of the payments even though it could not fully support this administrative position under then-current law. I.R.S. CCA 200114044; I.R.S. CCA 200114045. A limited exception was made for amounts "received for luxuries or for living expenses of an individual who has abandoned efforts to re-occupy a dwelling comparable to the one whose occupancy or use was denied by the fire." I.R.S. CCA 200114045.

Individuals would be able to exclude assistance from a charitable organization or another individual as gifts, so long as the transfer proceeded out of charitable impulses and without the imposition of quid pro quo conditions. *See* I.R.C. § 102(a) (2006); *Comm'r v. Duberstein*, 363 U.S. 278, 285 (1960).

²⁸⁹ In particular, a revenue ruling permitting employees to exclude disaster relief from employers was particularly problematic because it took the position that such transfers were not income because "[t]he objective of the corporation is to try to place the employees in the same economic position, or as near to it as possible, which they had before the casualty." Rev. Rul. 131, 1953-2 C.B. 112, 113 (1953), made obsolete by I.R.C. § 102(c) & I.R.C. § 139. The ruling did not, however, allow the employees to increase basis in damaged property. *Id.* at 113-14. The revenue ruling was issued prior to the Supreme Court's determination that income consisted of "undeniable accessions to wealth, clearly realized, and over which the taxpayers have complete dominion." *Comm'r v. Glenshaw Glass*, 348 U.S. 426, 431 (1955). But how casualty events should be treated even given an expansive definition of income remains a matter of debate. *See* Jeffrey H. Kahn, *Personal Deductions—A Tax "Ideal" or Just Another "Deal"?*, 2002 L. REV. M.S.U.-D.C.L. 1, 37-40 (2002) (arguing that casualty and theft loss deductions should not be treated as departures from economic income and should not be treated as tax expenditures by the Joint Committee on Taxation). *See also* Boris I. Bittker, *Income Tax Deductions, Credits, and Subsidies for Personal Expenditures*, 16 J.L. & ECON. 193, 198 (1973) (arguing that an insistence that there is only one way to view casualty losses in terms of an income definition is "sheer dogmatism").

More problematic for the validity of the ruling was the 1986 enactment of a rule prohibiting an exclusion from gross income for "any amount transferred by or for an employer to, or for the benefit of, an employee." I.R.C. § 102(c) (2006); Tax Reform Act of 1986, Pub. L. No. 99-514, 100 Stat. 2110 § 122(b) (1986). *See also* Rev. Rul. 2003-12, 2003-1 C.B. 283 ("[T]he payments made by the employer described in Rev. Rul. 131 do not qualify as gifts under § 102 and are not excluded from the employees' gross income under the general welfare exclusion.").

In the aftermath of September 11, 2001, Congress amended the Internal Revenue Code²⁹⁰ to codify partially the Service's administrative positions with respect to non-insurance disaster transfers without supplanting the exclusion for governmental general welfare transfers.²⁹¹ Currently, the Code provides an exclusion from gross income for a "qualified disaster relief payment"²⁹² (relief payment) and a "qualified disaster mitigation payment"²⁹³ (mitigation payment). Relief payments are tied to the immediate aftermath of a disaster while mitigation payments are grants to be used for improvements that will lessen the extent of future losses.²⁹⁴ With respect to either type of payment, the Code provides that if an excludible payment is received, the individual may not use the excluded funds to take a further deduction or credit.²⁹⁵ In other words, taxpayers may not obtain two tax benefits for the same dollars.

1. Qualified Disaster Relief Payments

In order to be excluded from income as a qualified disaster relief payment, the payment must be for "reasonable and necessary personal, family, living, or funeral expenses" or "reasonable and necessary expenses

Legislation enacted in the aftermath of September 11, 2001, clarifies that payments made in connection with certain types of disasters are not gross income, regardless of source of payment (other than insurance payments). I.R.C. § 139. *See also* Cords, *supra* note 224, at 442 (discussing difficulty of excluding payments from employer to victim-employee "because they did not easily fit within the definition of a gift").

²⁹⁰ Victims of Terrorism Tax Relief Act of 2001, Pub. L. No. 107-134 § 111, 115 Stat. 2427, 2432 (2001).

²⁹¹ Rev. Rul. 2003-12, 2003-1 C.B. 283 (explaining that § 139(b)(4) "codifies (but does not supplant) the administrative general welfare exclusion").

²⁹² I.R.C. § 139(b) (2006).

²⁹³ I.R.C. § 139(g).

²⁹⁴ *See* JOINT COMMITTEE ON TAXATION, GENERAL EXPLANATION OF TAX LEGISLATION ENACTED IN THE 109TH CONGRESS, JCS- 1-07, at 6 *Report* JCS-1-07 (explaining that mitigation payments are "grant[ed] to mitigate potential damage from future hazards" whereas relief payments ally to "certain amounts received by individuals as a result of a disaster that has occurred"); *see* 42 U.S.C. § 4011(b)(4) (listing mitigation programs and including properties covered by such programs in the NFIP); 44 C.F.R. Parts 78-80 (flood mitigation assistance & grants; property acquisition & relocation for open space).

²⁹⁵ I.R.C. § 139(h) ("[N]o deduction or credit shall be allowed . . . for, or by reason of, any expenditure to the extent of the amount excluded under [section 139] with respect to such expenditure.").

incurred for the repair or rehabilitation of a personal residence or repair or replacement of its contents.”²⁹⁶ FEMA temporary assistance grants would be excludible under this authority.²⁹⁷ A qualifying relief payment does not, however, include payments received under an insurance contract or compensation for costs that have already been covered by an insurance contract.²⁹⁸ Payments made under a flood insurance contract are, of course,

²⁹⁶ I.R.C. §§ 139(c)(1)-(2) (2006). In addition, non-governmental payment must be made in connection with a qualified disaster, which includes a disaster resulting from a “terroristic or military action” or a federally declared disaster. These items “terroristic or military action” are in turn defined in Code § 692(c)(2) and includes “any terroristic activity which a preponderance of the evidence indicates was directed against the United States or any of its allies” and “any military action involving the Armed Forces of the United States and resulting from violence or aggression against the United States or any of its allies (or threat thereof).” I.R.C. § 692(c)(2). Code section 692(c)(2) goes on to specify that “‘military action’ does not include training exercises.” *Id.* Code section 139 also has provisions relating to common carrier disasters (e.g., airline crashes). *See* I.R.C. §§ 139(b)(3), 139(c)(3).

A federally declared disaster “means any disaster subsequently determined by the President of the United States to warrant assistance by the Federal Government under the Robert T. Stafford Disaster Relief and Emergency Assistance Act.” I.R.C. § 165(h)(3)(C). The definition of “federally declared disaster” is in a Code section whose primary effect (an increase in the standard deduction) only applies to disasters occurring before January 1, 2010. Nevertheless, the definition itself has not expired. In any case, prior to 2008, section 139 was tied instead to a “Presidentially declared disaster,” which had virtually the same meaning as the current “federally declared disaster.” Tax Extenders and Alternative Minimum Tax Relief Act of 2008, Pub. L. No. 110-343, 122 Stat. 3765, 3922 Division C, § 706(a)(2)(D)(iv) (2008). In order for a disaster to be considered presidentially declared, the disaster, “with respect to the area in which the property is located, resulted in a subsequent determination by the President that such area warrants assistance by the Federal Government under the Robert T. Stafford Disaster Relief and Emergency Assistance Act.” I.R.C. § 1033(h)(3) (2006).

Payments from government sources, whether federal, state, or local, are excluded if made “in connection with a qualified disaster in order to promote the general welfare.” IRC 139(b)(4). With respect to governmental payments, qualified disaster is defined more broadly to include “a disaster which is determined by an applicable Federal, State, or local authority . . . to warrant assistance from the Federal, State, or local government or agency or instrumentality thereof.” I.R.C. § 139(c)(4).

²⁹⁷ *See* Lipman, *supra* note 183, at 962-71.

²⁹⁸ Other than these restrictions related to insurance coverage, the statute does not require that non-governmental payments must be from a particular source. For

ultimately made by the federal government, but the statute does not contain an exception for flood insurance in its requirement that the provision only applies to payments “not otherwise compensated by insurance.” Since the NFIP is generally treated as insurance for other purposes,²⁹⁹ and since the Service has apparently not issued guidance on this issue,³⁰⁰ payouts under the NFIP should be handled under the tax provisions relating to insurance recoveries rather than under the exclusion for governmental disaster relief payments.

For flood losses that occur outside the context of a federally declared disaster, individuals would be able to exclude transfers from individuals, charitable organizations, and government by arguing that these transfers are gifts (if from individuals or charities) or are general welfare transfers (if from government).³⁰¹ Thus, the main difference between treatment of flood losses occurring in federally declared disasters and other flood losses is that transfers by employers and employer-operated foundations would be subject to much greater scrutiny and would most likely be taxable as a matter of positive law (whether the Service would

example, the Service has confirmed that transfers from employers to employees may be excluded from income by the employees, so long as the requirements of the Code are met. Rev. Rul. 2003-12, 2003-1 C.B. 283 (holding that even though employer transfers to employees do not qualify as gifts or as excludible general welfare, they may qualify for the section 139 exclusion if the other conditions are met). *See supra* note 158 (discussing the problem of employer temporary assistance payments).

²⁹⁹ *See supra* Part II.

³⁰⁰ In 2000, the National Park Service caused a fire that destroyed more than 200 residences in New Mexico. The Service’s Office of Chief Counsel issued informal letters advising the exclusion of the FEMA payments made both to provide relief for the disaster and to settle any claims an individual might have against the federal government for the disaster. I.R.S. CCA 200114044; CCA 200114045. In addition, the Office of Chief Counsel further advised that FEMA reimbursements for NFIP premiums were excludible to the extent the fire caused taxpayers to need to purchase flood insurance as a result of the fire. I.R.S. CCA 200114046. The Chief Counsel’s Office provided little analysis to support its “belie[f] that under the unique circumstances . . . the government’s reimbursements of flood insurance premiums need not be treated as gain.” I.R.S. CCA 200114046. In any case, none of the Chief Counsel Advice memoranda dealt with flood insurance contract payments made to compensate for flood loss, and the letters also pre-date Section 139’s exclusion for qualified relief payments.

³⁰¹ *See* Rev. Rul. 2003-12, 2003-1 C.B. 283 (discussing treatment of government, charitable, and employer transfers in context of presidentially declared disaster).

pursue such transfers at the individual flood victim level is a different matter).

The proposed flood loss security program with its mandate may decrease the extent to which individuals receive aftermath aid from other sources.³⁰² Post-disaster assistance seems unlikely, however, to dwindle altogether, and in the case of high-profile events is still likely to be significant. This Article proposes that the exclusion for post-disaster assistance should continue given the possibility of unexpected needs and the administrative difficulty of enforcing an inclusion at a time of crisis. The exclusion should, however, be made more generous so as to apply with respect to any flood loss without the need for a federally declared disaster.

The current exclusion is allowed only to the extent amounts are not already covered by insurance.³⁰³ As an enhancement to the social safety net aspects of the proposed program and for administrative convenience, this provision could be lifted to the extent of the minimum required coverage for personal property.³⁰⁴ For example, if the minimum required coverage for personal possessions were \$15,000, individuals could receive a matching amount from non-insurance sources income-tax free even if there is some coverage duplication. It would, however, also be advisable to put a cap on the amount that could be excluded if received from non-government sources, especially employers. This cap could be set to match the personal property coverage maximum and would be added to prevent the problem of disguised compensation but also to avoid the possible creation of a shadow, government-subsidized system for higher income individuals.³⁰⁵ Of course, transfers from family and friends that exceed such a maximum amount would still potentially be excluded from income under the general provision for gifts.³⁰⁶

³⁰² See Levmore & Logue, *supra* note 4, at 280 (speculating that if private insurance covers disaster losses “there is apt to be less sympathy and therefore a lower probability of public or charitable relief”).

³⁰³ See *supra* Part IV.A.1.

³⁰⁴ This assumes that the proposed program is mandatory at all risk levels. See *supra* Part IV.B. If the program is made opt-out for certain categories of risks, the exclusion for non-insurance assistance should be largely disallowed as to those who choose to opt-out. This disallowance could help discourage individuals from opting out. The general gift provision of Code section 102 would still be available.

³⁰⁵ To the extent high-end private flood insurance is (or becomes) available, payments under the contract would be governed by general provisions applicable to insurance reimbursements. See *infra* Part V.B.

³⁰⁶ I.R.C. § 102 (2006).

2. Qualified Disaster Mitigation Payments.

In addition to aftermath aid, Congress has instituted grant programs aimed at lessening future flood damage.³⁰⁷ For example, homeowners may apply for grants to elevate a home.³⁰⁸ In an informal memorandum, the Service's Office of the Chief Counsel advised that such payments were taxable because they were for the mitigation of future disasters and thus were not the type of relief payments excluded by either the Code or the administrative general welfare exclusion.³⁰⁹ Congress acted in 2005 to change this result and provided a retroactive exclusion for these types of payments.³¹⁰ In order to qualify for the statutory exclusion, however, the payments may not be for the sale of the property.³¹¹ If the grant is in substance the purchase of a property, then payments are not excluded and would instead generate gain or loss according to the difference between the payment and the individual's tax investment (i.e., the individual's adjusted basis) in her property.³¹² Taxpayers would be able to defer recognition of any resulting gain through purchase of qualifying replacement property,³¹³

³⁰⁷ See I.R.S. CCA 200431012, 2004 WL 1701305 (IRC CCA) (describing mitigation grants authorized by the Robert T. Stafford Disaster Relief and Emergency Assistance Act and The National Flood Insurance Act); JOINT COMMITTEE ON TAXATION, GENERAL EXPLANATION OF TAX LEGISLATION ENACTED IN THE 109TH CONGRESS, JCS- 1-07.

NFIP flood insurance contract payments would not be excluded as qualified disaster mitigation payments since the contract payments are reimbursements for losses that have already occurred and are not made to lessen future losses.

³⁰⁸ See *supra* note 292 (mitigation programs).

³⁰⁹ I.R.S. CCA 200431012 (advising that the mitigation payments were not excludable under Code sections 102, 139, or 1033 or through administrative practice regarding general welfare or government-created property rights).

³¹⁰ I.R.C. § 139(g)(3) (2006); Public Law 109-7, § 1(a)(1) (2005). Any hazard mitigation payment used with respect to property may not also increase the basis in that property. I.R.C. § 139(g)(3); Pub. L. No. 109-7, 119 Stat. 21 § 1(a)(1) (2005). Any hazard mitigation payment used with respect to property may not also increase the basis in that property.

³¹¹ I.R.C. § 139(g)(2).

³¹² I.R.C. § 1001.

³¹³ Code section 1033(k) provides that section 1033 is available for these types of sales even though these programs are voluntary. I.R.C. § 1033(k). See *infra* Part IV.B.2 (describing Code section 1033).

but any loss would apparently be nondeductible if the payment related to a residence or other personal-use property.³¹⁴

The exclusion for hazard mitigation payments should remain in place³¹⁵ in order to continue to encourage steps that lessen the costs of flood loss. Expansion of the benefits of such steps should also be explored. For example, relocation programs that are the equivalent of a sale could provide for a loss deduction, if any tax loss results.³¹⁶

³¹⁴ If the sale relates to a personal residence, the loss would be a nondeductible personal loss under the Code. Individuals may deduct casualty and theft losses even if the underlying asset is a personal-use asset, but any loss generated by the type of sale described in Code section 139(g) would not qualify. The programs described in Code section 139(g) are voluntary hazard mitigation programs, so there is no involuntary taking. *See* CCA 200431012; Joint Committee Report, JCA-1-07.

Code section 165(k) does allow taxpayers to take a casualty loss deduction if a taxpayer is ordered by a governmental entity to demolish or relocate a residence because it has been rendered unsafe as the result of a federally declared disaster and the order to demolish occurs not later than the 120th day after the federal disaster declaration. I.R.C. § 165(k) (2006). Section 165(k) would not apply with respect to the voluntary hazard mitigation programs currently offered under the Stafford Act and the Flood Insurance Act. In the absence of section 165(k), it is less clear whether a government action such as an ordered demolition of an unsafe building would qualify as a casualty event. *See, e.g.,* Powers v. Commissioner, 36 T.C. 1191 (1961) (no casualty loss deduction allowed for impounding of car by East Berlin authorities); Washington v. Comm’r, T.C. Memo 1990-386 (losses arising out of a court-ordered eviction were not casualty losses). *Compare* I.R.C. § 280B (disallowing deduction for demolition costs). Eminent domain actions by federal, state, or local government require, of course, payment of just compensation. *See* Stop the Beach Renourishment, Inc., v. Fl. Dept. Environment’l Protection, 130 S. Ct. 2592, 2601-02 (2010) (general discussion of Takings Clause of U.S. Constitution). Because of the compensation element, section 1033 rather than section 165 would almost certainly be the applicable provision. Section 1033 is discussed *infra* Part V.B.2.

³¹⁵ *See supra* Part IV.A.2.

³¹⁶ This could be accomplished either by treating the loss as a casualty loss or as an investment loss. If treated as a casualty loss, some of the current limitations on deductibility could be relaxed, as has been done in the past for certain types of casualty losses. *See infra* Part V.C. Because the sale would be of a personal residence, any loss would be nondeductible under current law, so legislation would also be required for such a loss to qualify as an investment loss. If treated as an investment loss, the loss would be capital and subject to various timing constraints on deductibility, which could also be relaxed. *See* I.R.C. §§ 1211(b), 1212(b) (limiting capital loss deduction to amount of capital gains plus \$3,000, with excess

Such tax enhancements should, however, require that the individual move to a low risk home and also require evidence that the home had a prior history of flooding. Such tax enhancements should, however, require that the individual move to a low risk home and also require evidence that the home had a prior history of flooding. A more generous tax treatment for sale-equivalent relocation may particularly be needed if coverage of structural components is structured to decrease and eventually disappear for repetitive loss to the same structure.³¹⁷ Hazard mitigation grants could also be used to support adjustments to the premium charged individuals. In addition to direct grant programs, tax credit programs could also be enacted to encourage home improvements that would decrease flood loss.³¹⁸

B. INSURANCE PROCEEDS

Current tax law divides casualty insurance payouts into two basic categories: payments for temporary living expenses and payments for property damage.

1. Temporary Assistance

A limited tax exclusion applies to insurance payments for temporary living expenses.³¹⁹ The Code exempts from tax insurance

carried forward). For individuals without capital gains, casualty loss treatment would provide the lower tax result because casualty losses yield an offset against ordinary income.

³¹⁷ See *supra* Part IV.B.

³¹⁸ Similar tax credit programs have been enacted with respect to energy efficient improvements. See, e.g., I.R.C. § 25D (2006). See also KUNREUTHER & MICHEL-KERJAN, *supra* note 6, at 264 (suggesting tax incentives as a way “to encourage residents to pursue mitigation measures is to provide tax incentives” and describing success of an earthquake loss mitigation program established by the city of Berkeley, California).

³¹⁹ I.R.C. § 123 (2006). In the absence of Code section 123, such temporary assistance transfers would be taxable. Treas. Reg. § 1.123-1(a)(5) (insurance payments for living expenses are includible in gross income except to the extent provided for in Code section 123). First, payments for temporary assistance made by an insurance company to an insured would never qualify as a tax-exempt gift since such temporary assistance would occur by operation of the insurance contract instead of out of charitable impulses. Second, since individuals have no deduction for personal consumption, a non-statutory exclusion of the insurance proceeds for such consumption would be problematic. I.R.C. § 262. At the same time, the

reimbursements for “living expenses incurred during such period for himself and members of his household resulting from the loss of use or occupancy” if the individual’s principal residence is damaged by casualty or if the individual is not able to enter his principal residence on government orders because of the threat of a casualty.³²⁰ “Principal residence” in this context “depends upon all the facts and circumstances in each case,” and includes also rented residences.³²¹ The exclusion applies only to living expenses and not to payments made for loss of income or for lost or damaged property.³²² A federally declared disaster is not a requirement, so this exclusion applies to any casualty event causing displacement from the principal residence. A taxpayer may exclude the insurance payment only to the extent the actual expenses incurred during the displacement exceed the normal expenses that would have been incurred but were avoided as a result of the casualty.³²³ As a result, the exclusion applies only to duplicative and increased living expenses.³²⁴

dividing line between insurance premiums and pre-payments for services is not always clear. For example, purchasers of AAA undoubtedly view roadside assistance as services for which they have already made payments. *Cf.* *Am. Auto. Ass’n v. United States*, 367 U.S. 687, 689 (1961); *Auto. Club of Mich. v. Comm’r*, 353 U.S. 180, 180 (1957) (membership dues included in income upon receipt). Discussion of the line between prepayment for services and insurance is beyond the scope of this Article.

³²⁰ I.R.C. § 123(a); Treas. Reg. § 1.123-1(a).

³²¹ Treas. Reg. § 1.123-1(c). Omitted from this definition of “principal residence” is a link to Code section 121, which provides an exclusion for gains realized on the sale of a principal residence. In any case, the principal residence definition in the section 123 regulations is consistent with, if not as nuanced as, that contained in the section 121 regulations. *See* Treas. Reg. § 1.121-1(b). *See supra* Part IV.B (discussing principal residence concept).

³²² Treas. Reg. § 1.123-1(a)(3).

³²³ Treas. Reg. § 1.123-1(b)(1). The regulations also require that payments must be traceable to reimbursement for living expenses under the insurance contract. Thus, if an insured receives a payment on account of lost rental income and uses it for duplicative living expenses, the payment will not be excluded under this provision. The regulations contain ratios for determining the extent to which an insurance reimbursement is for living expenses if there is blanket coverage rather than identifiable living expense coverage. Treas. Reg. § 1.123-1(a)(4).

³²⁴ For example, if a family spends \$800 per month normally on food cooked in the residence but is now forced to spend \$1,200 on restaurant meals but spends nothing on cooking food, only a maximum of \$400 could be excluded for increased food costs. All the living expenses are considered in the aggregate, so this \$400 increase might be offset by decreases elsewhere—for example, by a

This limitation is not lifted even if the triggering event is a federally declared disaster. In at least one instance, the Service's Office of Chief Counsel has, however, advised against implementing the limitation and instead advised field agents to apply a blanket exclusion for all living expense reimbursements other than those for "luxuries or for living expenses of an individual who has abandoned efforts to re-occupy a [comparable] dwelling."³²⁵ The Chief Counsel Advice memorandum may not be used as precedent and was given in response to a disaster triggered by the actions of the National Park Service.³²⁶ Still, the memorandum perhaps provides some indication of the relative zeal with which the Service will audit those claiming exclusions for insurance coverage of temporary living expenses — particularly if the displacement occurs in the context of a large-scale disaster.

Coverage under the NFIP does not currently cover temporary living expenses,³²⁷ but the proposed program would provide such coverage and the current exclusion would thereby become applicable. The limitation relating to the need for duplicative costs should be lifted with respect to flood program payments for the same reasons that non-insurance amounts should be excluded up to a certain point: to further safety net goals and to reduce administrative complexity.³²⁸ An exception for luxuries, as suggested in the Chief Counsel Advice described above, should not be necessary because the amount of coverage for temporary living expenses would be statutorily capped at an amount tied to meeting basic needs.

2. Property Loss Reimbursement

It may seem counterintuitive that a catastrophe could give rise to a tax liability, but such is the case if reimbursements for property exceed the taxpayer's investment in the property. For example, if an individual purchased a painting for \$100,000 many years ago and receives \$300,000

reduction in commuting costs. Treas. Reg. § 1.123-1(b)(4) Ex. 1. Professor Lipman has noted that even though these provisions are strict on their face "in practice they may have little application. Homeowner's insurance coverage generally only reimburses a homeowner for additional living expenses, which is defined consistently with the exclusion provision." Lipman, *supra* note 181, at 984-85.

³²⁵ I.R.S. TECH. ADV. MEM. 200114045 (Mar. 29, 2001).

³²⁶ *Id.*

³²⁷ See NAT'L FLOOD INS. PROGRAM, *National Flood Insurance Program Summary of Coverage*, http://www.floodsmart.gov/floodsmart/pages/residential_coverage/whats_covered.jsp.

³²⁸ See *supra* Part V.A.1.

in insurance proceeds for the painting, the individual would recognize \$200,000 of casualty gain. The Code permits taxpayers to elect to defer paying taxes on such casualty gains by purchasing replacement property.³²⁹ The replacement property must be “similar or related in service or use” to the original, destroyed property.³³⁰ The gain is deferred rather than completely excluded by treating the taxpayer as though his investment in the replacement property is carried over from the destroyed property.³³¹

As discussed above, property purchased as part of a hazard mitigation program is eligible for deferral of any gain on the sale.³³² A special rule also applies to principal residences that are “compulsorily or involuntarily converted as a result of a Presidentially declared disaster”³³³; taxpayers receive a full exclusion for insurance proceeds received for unscheduled property without the need to purchase replacement property.³³⁴ The extent to which household contents are treated as unscheduled property will depend on the particular insurance contract. In general, only assets of relatively high value (e.g., jewelry, artwork) will be separately scheduled, and all other household property will be treated as a single asset.³³⁵ The

³²⁹ I.R.C. § 1033 (2006).

³³⁰ I.R.C. §§ 1033(a)(1)-(2). *See* BORIS I. BITTKER ET AL., *FEDERAL INCOME TAXATION OF INDIVIDUALS* ¶ 30.03[3]-[5] (3d ed. & 2010 cumulative supplement) (discussing this requirement).

³³¹ I.R.C. § 1033(b).

³³² *See supra* Part V.A.2.

³³³ I.R.C. § 1033(h)(1) (2006). The definition for principal residence is tied to section 121 through a cross-reference, although a home may qualify even if rented rather than owned. I.R.C. § 1033(h)(4). *See supra* Part IV.B.

Section 121, which provides a generous exclusion for gain on the sale of a principal residence if various eligibility requirements are met, may also be available. I.R.C. § 121(d)(5) (amount realized on involuntary conversion of principal residence is reduced by amount of section 121 exclusion).

Because sales pursuant to a hazard mitigation program are not in response to a federally declared disaster but are instead aimed at lessening future losses, presumably Code section 1033(h)(1) does not apply to mitigation sales. A special rule expanding the scope of qualifying replacement property also applies to trade, business, or investment property converted as a result of a federally declared disaster. I.R.C. § 1033(h)(2).

³³⁴ I.R.C. § 1033(h)(1)(A)(i) (2006). *See also* Rev. Rul. 95-22, 1995-1 C.B. 145.

³³⁵ *See* Rev. Rul. 95-22 1995-1 C.B. 145 (containing example situation in which general household furnishings were unscheduled while jewelry and sterling silverware were separately scheduled). The IRS Chief Counsel has advised in an informal memorandum that the exclusion will apply even to property not kept at

principal residence (if owned) and any separately scheduled contents are treated as a single asset for purposes of calculating gain and applying the deferral provision.³³⁶ In addition, the replacement property may be anything “which is similar or related in service or use to the residence so converted”— or its contents.³³⁷

Thus, for example, consider a taxpayer whose rented residence and its contents are destroyed in a federally declared disaster and who receives \$40,000 of insurance proceeds for unscheduled household items and \$10,000 for scheduled jewelry. All \$40,000 of the proceeds received for the unscheduled household items will be excluded from income even if no replacement property is purchased.³³⁸ If the taxpayer originally purchased the jewelry for \$8,000, the \$2,000 gain arising from the \$10,000 insurance payment can be deferred through purchase of \$10,000 of replacement property. The replacement property may be jewelry but it may also be any household-related item — e.g., linens, dishes, furniture.³³⁹

These provisions should be expanded to cover all flood-related reimbursement received under the proposed program. Coverage for personal possessions would be treated as payment for unscheduled property and thus any gain would be excluded from income. The exclusion for gain resulting from reimbursement for unscheduled property is not, however, as generous as it appears on its face. That is because it will be relatively rare for a taxpayer to have a gain for typical, unscheduled household furnishings, which generally go down in value after purchase. If insurance

the principal residence so long as the property was covered by the contract and was lost in the federally declared disaster (e.g., property in a car). I.R.S. TECH. ADV. MEM. 200114046 (Apr. 2, 2001).

³³⁶ I.R.C. § 1033(h)(1)(A)(ii)(I) (2006). Presumably the section 121 adjustment would apply to this aggregated asset, assuming the other qualifications of section 121 are met. I.R.C. § 121(d)(5). The adjustment reduces the amount realized from insurance by the amount excluded under section 121. Thus, for example, if \$700,000 is received under an insurance contract for a qualifying principal residence by a couple filing a joint return, section 1033 is applied as though only \$200,000 of insurance proceeds were received. I.R.C. § 121(a)-(b), (d)(5).

³³⁷ I.R.C. § 1033(h)(1)(A)(ii)(II). The time for purchasing the replacement property is also increased from two to four years. I.R.C. § 1033(h)(1)(B).

³³⁸ Rev. Rul. 95-22 1995-1 C.B. 145 (no gain recognized “upon the receipt of insurance proceeds for unscheduled contents destroyed in such a disaster, regardless of the use to which the taxpayer puts those proceeds”).

³³⁹ See Rev. Rul. 95-22 1995-1 C.B. 145 (“[A]ny type of replacement contents (whether separately scheduled or unscheduled)” qualifies as replacement property for separately scheduled contents).

proceeds are insufficient to reimburse a taxpayer for his investment in the property, the taxpayer will have a loss. For example, if a taxpayer purchased an asset for \$10,000 but receives only a \$6,000 insurance recovery, the taxpayer has \$4,000 tax loss. Whether such a loss is or should be deductible is considered below.

C. UNREIMBURSED LOSSES

Losses that arise from the disposition of personal-use assets are generally nondeductible.³⁴⁰ Taxpayers may, however, take a limited deduction if the loss is caused by a casualty event or theft and is not compensated for by insurance or through some other means.³⁴¹ The calculation of the casualty loss deduction is complex and requires a series of steps, each of which potentially serves to limit the size of a deduction.³⁴²

³⁴⁰ I.R.C. § 165(c)(3) (2006).

³⁴¹ Whether an event constitutes a casualty event or theft loss is itself a difficult issue to resolve. *See* BITTKER ET AL., *supra* note 330, at ¶¶ 24.02-.03; BURKE & FRIEL, *supra* note 288, at 530-33.

³⁴² First, the amount of the loss is limited to the lesser of the taxpayer's investment in the asset or the decline in value of the asset. Treas. Reg. § 1.165-7(b)(1). Thus, for example, if a taxpayer spent \$10,000 for an asset but the asset was worth only \$7,000 when it was destroyed in a flood, the amount of the potential casualty loss would be limited to \$7,000. Second, the loss is reduced by the extent to which an individual receives compensation for the loss (whether through insurance or otherwise). I.R.C. § 165(a) (2006); Treas. Reg. § 1.165-1(d). If the taxpayer received \$6,000 in insurance proceeds, his potential casualty loss deduction would be further reduced to \$1,000.

Third, the Code disallows the first \$100 of casualty loss stemming from a casualty event or theft. In the case of the example, the taxpayer's potential casualty loss deduction would be reduced to \$900. I.R.C. § 165(h)(1). This \$100 amount was temporarily raised to \$500 during 2009. Tax Extenders and Alternative Minimum Tax Relief Act of 2008, P.L. 110-343, Division C, § 706(c), 122 Stat. 3921-3923.

Fourth, the sustained casualty losses for the year are aggregated and applied first against the aggregate of any casualty gains for the year. I.R.C. § 165(h)(2)(A)(i). For example, if the taxpayer has a \$400 gain resulting from an unrelated theft, only \$500 would remain as the net casualty loss. If the taxpayer elects to defer the gain under section 1033, then the casualty gain is not included in this netting calculation.

Finally, once the net casualty loss amount is determined, casualty losses are only deductible to the extent they exceed 10% of a taxpayer's adjusted gross income. I.R.C. § 165(h)(2)(A)(ii). Thus, if a taxpayer has a \$500 potential casualty

For example, losses are only deductible to the extent they exceed ten percent of adjusted gross income; thus, only relatively large casualty losses are deductible as a practical matter. In addition, because the casualty loss deduction is an itemized deduction, the value of the deduction will increase with higher rate brackets and also will not be available to those who use the standard deduction rather than itemize.³⁴³ The ten-percent-of-adjusted-gross-income threshold may, however, still place the deduction out of reach for individuals with high taxable income. During 2008 and 2009,³⁴⁴ casualty losses caused by a federally declared disaster were less limited as to amount³⁴⁵ and could also be used without the need to itemize.³⁴⁶

The deduction for casualty losses should be significantly limited, if not eliminated, for flood losses.³⁴⁷ The deduction creates a shadow system

loss deduction, and the taxpayer's adjusted gross income is \$5,000 or more, no casualty loss deduction will be permitted.

Even if an amount of net casualty loss remained over the 10% floor, the taxpayer would only get the benefit of the deduction by electing to itemize his deductions rather than taking the standard deduction. I.R.C. § 63(b)-(c) (2006). In 2011, the standard deduction for a taxpayer filing as single will be \$5,800; for head of household, \$8,500; and for married filing joint, \$11,600.

³⁴³ For example, at the margin, a \$1,000 deduction is worth \$350 to someone in a 35% rate bracket but only \$200 to someone in a 20% bracket.

³⁴⁴ I.R.C. § 165(h)(3)(B)(i)(I); Tax Extenders and Alternative Minimum Tax Relief Act of 2008, P.L. 110-343, Division C, § 706122 Stat. 3921-3923. A similar measure was also in place in the immediate aftermath of Hurricane Katrina. Katrina Emergency Tax Relief Act of 2005, Pub.L. No. 109-73, § 402. *See* Aprill & Schmalbeck, *supra* note 219, at 59-60 (discussing temporary change to section 165 for 2005 hurricane season losses).

Legislative proposals have been made to extend or make permanent the standard deduction increase for net disaster losses. *See* H.R. 5273, 111th Cong. (2010); H.R. 4213, 111th Cong. (2010); and H.R. 4052, 111th Cong. (2009).

³⁴⁵ They were not limited by the ten-percent floor. I.R.C. § 165(h)(3)(A) (2006). Currently, a taxpayer who experiences casualty losses as the result of a federally declared disaster may elect to deduct the casualty losses on the tax return for the year preceding the disaster. I.R.C. § 165(i). *See also supra* note 314 (discussing § 165(k)). Personal casualty losses are not carried forward so if there is not enough income to soak up the casualty loss, any tax benefit to be obtained from the deduction would be lost. A taxpayer whose income fell as a result of the federally declared disaster could use the election to move the deduction to a year in which the taxpayer had income against which to offset the deduction.

³⁴⁶ The Code section did so by increasing the standard deduction. I.R.C. § 63(c)(2) (2006).

³⁴⁷ Sorting costs would need to be taken into account. *See* Levmore & Logue, *supra* note 4, at 321-22 (discussing sorting costs that would result from

of government reimbursement for flood loss, but one that offers patchy, difficult-to-understand coverage.³⁴⁸ Individuals may overestimate its benefits, which may in turn contribute to less care with respect to purchase decisions.³⁴⁹ For example, a relatively wealthy individual may purchase a \$400,000 beach house, purchase the maximum NFIP policy on the home, and assume that the remaining \$150,000 loss would be deductible should the home be destroyed in a flood. But if the individual's adjusted gross income were \$750,000, only approximately half of the \$150,000 loss would be deductible.³⁵⁰ As discussed in Part III, flood loss is difficult for individuals to conceptualize, so the effect of this shadow system may be relatively small. At the same time, individuals who itemize and who have the ability to purchase that second vacation home may be particularly tempted to believe that there is little personal downside to such a purchase given the combination of the NFIP and the casualty loss deduction.

Eliminating the casualty loss deduction for flood-related costs altogether would, however, likely be politically impractical. One possible compromise would be to allow the loss deduction but only for flood losses occurring at the principal residence and only for a limited dollar amount. The deduction should be accessible even those who do not itemize, and the ten-percent floor and other limitations should be lifted to provide greater certainty about the amount to be deducted. For example, the deduction could be limited to the loss in excess of the purchased coverage and up to an additional \$50,000 for personal property and \$125,000 for structural damage (these amounts are one-half the proposed coverage maximums). If, however, eligible coverage has been phased out for a structure because of repetitive claims, the loss deduction should be commensurately reduced. Elimination of the casualty loss deduction as to second homes and to a portion of the cost of more expensive homes would still be controversial,

government-sponsored crime insurance that did not cover terrorism). The same casualty may trigger flood loss, windstorm loss, or possibly even fire loss. In many cases, the presence of private insurance could simplify the inquiry. For example, the Code could treat private insurance recoveries as being for losses other than flood and assigning any residual loss as flood loss.

³⁴⁸ See Louis Kaplow, *The Income Tax as Insurance: The Casualty Loss and Medical Expense Deductions and The Exclusion of Medical Insurance Premiums*, 79 CAL. L.R. 1485 (1991). See also *supra* Part III.B (discussing difficulties individuals face in processing flood loss probability and calculating assistance).

³⁴⁹ See Kaplow, *supra* note 10.

³⁵⁰ Ten percent of \$750,000 is \$75,000. The deductible amount would be \$74,900 -- \$150,000 minus \$100 minus \$75,000.

but the program goals of providing a social safety net while limiting repetitive loss would be better served.

VI. CONCLUSION

Use of tax system components to implement social programs should not be lightly undertaken. In the case of flood loss mitigation, the government already plays a central function through both the NFIP and various tax provisions already in place. Moving flood loss coverage under the umbrella of the tax law could yield significant benefits, including increased program efficiencies and better tools for balancing competing land use goals. Most importantly, a national flood security system would be a means of providing the least fortunate with a safety net when (not if) the next unimaginable flood occurs.

A CONCURRENT MESS AND A CALL FOR CLARITY IN FIRST-PARTY PROPERTY INSURANCE COVERAGE ANALYSIS

MARK M. BELL*

This article clearly and plainly describes the genesis and history of the doctrine of "concurrent causation" and the development of anti-concurrent policy exclusions in first-party property insurance coverage cases. After describing this unique history, the article argues that it is time to create a new lexicon for "concurrent causation" issues and advocates for a new deliberate, categorical approach for addressing "concurrent causation" questions.

I. INTRODUCTION

Insurance coverage questions for “all-risk” policies are conceptually simple. If a peril is excluded, there is no coverage; if a peril is not excluded, there is coverage.¹ While this analysis seems conceptually simple, it becomes complicated in practice when multiple perils combine to cause a loss.

The complications are most acute when non-excluded, covered perils combine or operate in conjunction with excluded, non-covered perils to cause a loss. When covered and non-covered perils are connected to a

* Attorney with Waller, Lansden, Dortch & Davis, Nashville, Tennessee practicing complex commercial litigation with a focus on insurance and construction matters. I would like to thank one of the all-time property insurance greats, James Costner, for numerous discussions and feedback related to property insurance and concurrent causation. He has provided a lifetime worth of insights and his desire to retire will leave a notable absence in the insurance world. I would also like to thank David Rossmiller for thoughtful comments and feedback.

¹ This central tenant of insurance, however, is slowly being eroded as well. Mold exclusions are beginning to deny the result—mold—irrespective of what caused the mold. Historically, insurers excluded perils, but it seems that insurers are slowly beginning to exclude results. See J. Kent Holland Jr., *Mold from a Covered Concurrent Cause Still Excluded*, IRMI.COM (Nov. 2010), <http://www.irmi.com/expert/articles/2010/holland11-insurance-law-environmental.aspx>.

loss, it may be unclear from the policy whether the entire loss should be covered, whether the entire loss should be excluded, or whether the loss and resultant damages should be bifurcated to indemnify the insured for losses caused by covered perils while denying indemnity for losses caused by excluded perils.

Many courts and commentators refer to the process of multiple covered and excluded perils combining to cause a loss as “concurrent causation”.² It goes without saying, but nevertheless needs to be said, that the phrase “concurrent causation” presents a definitional problem. While the common definition of “concurrent” implies a degree of temporal simultaneity,³ courts and commentators have routinely used the term “concurrent” to refer to sequential chains of events;⁴ independent, unrelated events acting in conjunction;⁵ and even events that undoubtedly operated in succession.⁶ These types of events patently contradict the term “concurrent”; thereby turning “concurrent causation” into a definitional misnomer.

In addition to these definitional inconsistencies, courts have complicated the issues by developing a patchwork of interpretations of concurrent causation and relevant anti-concurrent causation policy exclusions.⁷ This resultant patchwork has operated to deprive policyholders

² William Conant Brewer, Jr., *Concurrent Causation in Insurance Contracts*, 59 MICH. L. REV. 1141, 1145 (1961). As will be discussed *infra*, the term “concurrent” has become imprecise, but it is used here as background. The term “concurrent” can have two distinct meanings. Today, concurrent either describes multi-cause losses operating or occurring at the same time or refers generically to a web of events having some interrelation among them.

³ See, e.g., AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE 383 (Joseph P. Pickett et al. eds., 4th ed. 2000) (“adj. Happening at the same time as something else.”); MERRIAM-WEBSTER’S COLLEGIATE DICTIONARY 239 (Frederick C. Mish et al. eds., 10th ed. 2001) (“adj. operating or occurring at the same time.”); WEBSTER’S NEW WORLD COLLEGE DICTIONARY 303 (Michael Agnes et al. eds., 4th ed. 2001) (“occurring at the same time; existing together.”).

⁴ *Churchill v. Factory Mut. Ins. Co.*, 234 F. Supp. 2d 1182, 1187 (W.D. Wash. 2002).

⁵ David P. Rossmiller, *Katrina in the Fifth Dimension: Hurricane Katrina Cases in the Fifth Circuit Court of Appeals*, NEW APPLEMAN ON INSURANCE: CURRENT CRITICAL ISSUES IN INSURANCE LAW 71, 77 (2008).

⁶ Peter Nash Swisher, *Causation Requirements in Tort and Insurance Law Practice: Demystifying Some Legal Causation “Riddles,”* 43 TORT TRIAL & INS. PRAC. L.J. 1, 4 (2007).

⁷ Douglas G. Houser, *The Rise and Fall of Concurrent Causation: Background and Current Trends Affecting Property Insurance Coverage*, 44 FED.

of their reasonable expectations and has prevented insurers from maintaining contract certainty when drafting insurance policies.⁸

Analyzing this patchwork of interpretations has left one commentator to explain, “[b]ecause causation as a theory or doctrine is so elusive, *inconsistent outcomes must be tolerated*.”⁹ As stated by this commentator, “[s]ometimes the different outcomes will turn on subtle factual distinctions, but sometimes the different outcomes will be based on an utterly irreconcilable view of policy text and principles of interpretation.”¹⁰

This article argues that inconsistent outcomes need not be tolerated, provides definitional clarification for the relevant elements of the concurrent causation phenomenon, and proposes a revised analytical framework to minimize the inconsistent outcomes. The article provides both a history of concurrent causation and a history of anti-concurrent policy exclusions. Using that history, the article proffers new definitions to address multi-cause losses,¹¹ and advocates for a more methodical, categorical analysis for addressing “concurrent causation” questions.¹²

II. BACKGROUND OF CONCURRENT CAUSATION

Courts have struggled with the question raised in the introduction on the best way to deal with losses caused by multiple perils. To address the issue, courts have typically employed one of four approaches to “concurrent” losses:¹³ the pro-policyholder approach, the pro-insurer approach, the dominant-cause approach, and the apportionment approach.¹⁴

OF INS. & CORP. COUNSEL Q. 3, 3 (1993).

⁸ See Michael E. Bragg, *Concurrent Causation and the Art of Policy Drafting: New Perils for Property Insurers*, 20 FORUM 385, 387-88 n.8 and accompanying text (1985).

⁹ ROBERT H. JERRY, II & DOUGLAS R. RICHMOND, UNDERSTANDING INSURANCE LAW 560 (4th ed. 2007) (emphasis added).

¹⁰ *Id.*

¹¹ As discussed *infra* p. 21-22, this article advocates that the phrase concurrent causation should be replaced with the more accurate term “multi-cause loss”.

¹² A flow chart setting out the interpretive mechanism is attached as Appendix A.

¹³ Peter Nash Swisher, *Insurance Causation Issues: The Legacy of Bird v. St. Paul Fire & Marine Ins. Co.*, 2 NEV. L.J. 351, 366-68 (2002).

¹⁴ Conventional scholarship does not typically refer to them by these names, but the names provided herein more adequately and easily describe the relevant categories. For the traditional names, see JERRY & RICHMOND, *supra* note 9, at

Interestingly, and generating further confusion, courts routinely refer to each approach as the “concurrent cause doctrine.”¹⁵

A. PRO-POLICYHOLDER APPROACH¹⁶

Under the pro-policyholder approach, if multiple perils combine to create a loss, the full amount of the loss is covered, so long as part of the loss was caused, even if insignificantly, by a covered cause of loss.¹⁷ This approach has also been referred to by courts as the “concurrent causation” doctrine or approach.¹⁸

The California Supreme Court in *State Farm v. Partridge* was one of the first courts to adopt this approach for **liability** policies.¹⁹ The

560-61.

¹⁵ See e.g., *Wallis v. United Services Auto Ass’n*, 2 S.W.3d 300, 302-03 (Tex. App. 1999) (applying the “concurrent cause doctrine” to the apportionment approach: “Texas recognizes the doctrine of concurrent causes. This doctrine provides that when, as in the instant case, covered and non-covered perils combine to create a loss, the insured is entitled to recover only that portion of the damage caused solely by the covered peril(s).”); *Allstate Ins. Co. v. Watts*, 811 S.W.2d 883, 886 (Tenn. Ct. App. 1991) (a third-party case applying the “concurrent cause doctrine” to the dominant cause approach: “The Court also opined that insurer was liable under ‘the concurrent cause doctrine’ which provides that coverage under a liability policy is equally available to an insured whenever an insured risk constitutes a concurrent proximate cause of the injury.”); *Farmers Ins. Exch. v. Adams*, 170 Cal. App. 3d 712 (Cal. Ct. App. 1985) (applying the “concurrent cause doctrine” to the pro-policyholder approach); *Wallach v. Rosenberg*, 527 So.2d 1386 (Fla. Dist. Ct. App. 1988) (same result in first-party context); *Phillips & Coplin*, *infra* note 87, at 33 (“A minority follows the doctrine of concurrent causation where coverage is afforded as long as a covered cause of loss contributes in a meaningful way to the insured’s damages.”). In this author’s opinion, the pro-policyholder use is the most accurate explanation of the “concurrent causation doctrine” because it was first and it spawned the anti-concurrent causation clauses proliferating property insurance policies today. See *infra* Part IV.

¹⁶ Throughout this article, when the term “post-*Partridge*” is used, it is in reference to the proliferation of the pro-policyholder approach.

¹⁷ For further discussion of the policy rationales supporting this approach, see *infra* note 114 and accompanying text.

¹⁸ See, e.g., *Allstate Ins. Co. v. Watts*, 811 S.W.2d 883, 886 (Tenn. 1991) (“[T]he ‘concurrent causation doctrine’ . . . provides that coverage under a liability policy is equally available to an insured whenever an insured risk constitutes a concurrent proximate cause of the injury.”).

¹⁹ *State Farm Mut. Auto. Ins. Co. v. Partridge*, 514 P.2d 123 (Cal. 1973). For further discussion, see *infra* Part III.B.

California Court of Appeals attempted to also adopt this approach for **property** policies in *Farmers Insurance Exchange v. Adams*.²⁰ In *Adams*, the court held that if third-party negligence (a covered loss) contributes in any respect to the loss, the entire loss is covered even if the efficient proximate cause of the loss would be excluded.²¹ While *Adams* was later overruled by the California Supreme Court, it demonstrates how courts analyze cases under the pro-policyholder approach.

In arguing for the pro-policyholder approach, courts reason that public policy militates in favor of the pro-policyholder approach.²² For instance, because ambiguities in insurance contracts of adhesion are generally interpreted strictly against the insurer and in favor of the insured, courts reason that when the loss is caused at least in part by a covered peril, the exclusion should be interpreted against the insurer.²³ Accordingly, under the pro-policyholder approach, when non-excluded perils and covered perils act in conjunction to cause the loss, the loss is covered.²⁴

B. PRO-INSURER APPROACH

The pro-insurer approach applies the opposite view of the pro-policyholder approach. Under the pro-insurer approach, if one of the causes of loss is excluded, the entire loss is excluded. While no domestic jurisdictions have entirely adopted this approach, British courts apply the pro-insurer approach with some uniformity.²⁵

²⁰ 170 Cal. App. 3d 712 (Cal. Ct. App. 1985).

²¹ *Id.* This case was later expressly rejected by the California Supreme Court in *Garvey v. State Farm Fire & Cas. Co.*, 770 P.2d 704, 711 (Cal. 1989).

²² *Premier Ins. Co. v. Welch*, 189 Cal. Rptr. 657, 660-62 (Cal. Ct. App. 1983).

²³ This is generally referred to as *contra proferentem* and applies in contracts of adhesion. For large commercial entities creating and negotiating manuscript policies, the same application may not apply. See JEFFREY W. STEMPEL, *STEMPEL ON INSURANCE CONTRACTS* § 4.11[F] (3d ed. 2006 & Supp. 2011) (“If the [manuscript policy] is essentially drafted by the policyholder, a weak version of *contra proferentem* should apply in reverse.”).

²⁴ This can also be attributed, at least in part, to California’s Insurance Code, which provides additional protections to policyholders. CAL. INS. CODE §§ 530, 532 (West 2011).

²⁵ See Erik S. Knutsen, *Confusion About Causation in Insurance: Solutions for Catastrophic Losses*, 61 ALA. L. REV. 957, 972-73 (2010). Some may argue that *Lydick v. Insurance Co. of North America*, 187 N.W.2d 602 (Neb. 1971) stands for the proposition that Nebraska follows this approach, but that case, and more recent cases in Nebraska indicate that the court was actually applying the dominant-cause

The British case, *Wayne Tank*, provides the quintessential example of this approach.²⁶ The Wayne Tank factory suffered a fire caused by two concurrent perils: “[F]ailure to install proper equipment (an excluded cause) and employee negligence in leaving the factory unattended (a covered cause).”²⁷ The *Wayne Tank* court held that even if the employee negligence was the predominant factor in the loss, the loss would still be excluded because the failure to install the proper equipment concurrently acted to cause the loss.²⁸ Thus, the entire loss was excluded because at least part of the loss was excluded.

It is unclear exactly why British courts have taken this approach, but perhaps it can be explained, at least in part, by the history of insurance in the United Kingdom and the remnants of a time when the insurer had less influence and control over the policy-making process. The roots of modern insurance date back to the United Kingdom and the shipping industry.²⁹ One of the first insurers, Lloyd’s of London, insured ships and their cargo and provides a fundamental building block for insurance interpretation in the United Kingdom.³⁰

In these pre-modern transactions, the insurer was at an information disadvantage to the shipper. The shipper had a better understanding of his skills and the unique challenges presented by his specific cargo, and also had significant control over his risks and potential losses. The insurer, conversely, was often at the mercy of the shipper and had to rely on the shipper providing truthful and accurate information to make its underwriting determinations. Because of this information asymmetry in favor of the shipper, Lloyd’s policies were often interpreted strictly against the shipper.³¹ For instance, if the shipper issued a warranty and that warranty was even partially breached, the entire loss was excluded.³² Accordingly, the British courts’ comparatively stern treatment of the policyholder may be rooted in this specific anachronism.

In the United States today, unlike the United Kingdom hundreds of years ago, insurance policies are contracts of adhesion and the justification

approach to the loss.

²⁶ *Id.* at 973 (citing *Wayne Tank & Pump Co. Ltd. v. Empr’r’s Liab. Assurance Co., Ltd.*, [1973] 3 W.L.R. 843 (Eng.)).

²⁷ *Id.* at 973.

²⁸ *Id.*

²⁹ See *JERRY & RICHMOND*, *supra* note 9, at 560-61 (citing *Shinrone Inc. v. Ins. Co. of N. Am.*, 570 F.2d 715 (8th Cir. 1978)).

³⁰ *Id.* at 561.

³¹ *Id.* at 749-50.

³² *Id.*

for strict interpretation of insurance policies against the insured no longer remains an attractive option. Today, insurers have the negotiating leverage, which is why United States' jurisdictions read ambiguities broadly against the insurer and read exclusions narrowly.³³ For this reason, American courts have been reluctant to follow the British, pro-insurer approach.

C. THE EFFICIENT PROXIMATE CAUSE RULE OR DOMINANT-CAUSE APPROACH

The dominant-cause approach attempts to strike a balance between the pro-policyholder and pro-insurer approaches and relies on equitable principles of fairness and the parties' reasonable expectations.³⁴ Under this approach, the court attempts to ascertain which cause, among the various concurrent causes of loss—or which link in the chain of events—was the most important, substantial, or responsible factor in the loss. This approach is also commonly referred to as the efficient proximate cause approach.³⁵

Shinrone Inc. v. Insurance Co. of North America demonstrates how courts apply the dominant-cause approach.³⁶ *Shinrone* involved a coverage dispute when cattle were killed during a storm with intense winds, damp snow, muddy land, and extremely cold temperatures.³⁷ The policy in question provided coverage for death by windstorm, but excluded death caused by “dampness of the atmosphere or extremes of temperature”.³⁸ The testimony in the case conflicted and experts concluded that the cattle died due to a combination of factors including wind, cold temperatures, snow, the size and age of the cattle, conditions of the land, and the lack of adequate wind protection.³⁹ Analyzing these factors, the jury determined that the windstorm was the most important or “efficient proximate cause” of the loss.⁴⁰ The jury reasoned that “extreme temperature” could not be the efficient proximate cause of the loss because without the wind, the cattle

³³ *Travelers Ins. Co. v. Aetna Cas. & Sur. Co.*, 491 S.W.2d 363, 366 (Tenn. 1973).

³⁴ See, e.g., *Allstate Ins. Co. v. Watts*, 811 S.W.2d 883, 886-88 (1991).

³⁵ Erik S. Knutsen, *Confusion About Causation in Insurance: Solutions for Catastrophic Losses*, 61 ALA. L. REV. 957, 974-75 (2010).

³⁶ *JERRY & RICHMOND*, *supra* note 9, at 561 (citing *Shinrone Inc. v. Ins. Co. of N. Am.*, 570 F.2d 715, 716 (8th Cir. 1978)).

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.*

⁴⁰ *Id.*

may have survived the extreme cold.⁴¹

The dominant cause approach has been adopted by the majority of U.S. jurisdictions because many consider it intuitively fair: decisions whether coverage should be afforded depend on which cause most significantly contributed to the loss.

D. THE APPORTIONMENT APPROACH

The final available approach is the apportionment approach, which Texas has adopted.⁴² Under the apportionment approach, the “insured must attempt to segregate the loss caused by the covered peril from the loss caused by the uncovered peril and secure a jury finding on the amount of damage attributable to the different causes.”⁴³ The approach follows traditional tort apportionment doctrines. As is the case with comparative negligence, there are two potential sub-approaches to the apportionment approach: pure apportionment and modified comparative apportionment.⁴⁴

Under a pure apportionment approach, the policyholder would receive the apportioned percentage of the damages caused by the covered losses. For instance, if 30% of the loss was caused by a covered peril, then the insured would receive 30% of the total value of the loss—or 30% of the policy limit if the limits were an issue.⁴⁵

Under a modified apportionment approach, the policyholder would receive the percentage of the loss so long as the efficient proximate cause was a covered peril.⁴⁶ Thus, if only 30% of the loss were caused by a covered cause of loss, the insured would not receive any recovery since the efficient proximate cause would have presumably been some other cause.

This approach inevitably leads to greater litigation and provides an

⁴¹ *Id.* This does impliedly reject the converse—that without the cold the wind could not have killed the cattle—but the jury did not address that issue.

⁴² *Wallis v. United Serv.s Auto Ass’n*, 2 S.W.3d 300 (Tex. App. 1999) (requiring insured to carry burden of proof for what portion of loss is covered). For an interesting discussion of Texas’s approach to anti-concurrent exclusions, see Comment, Amber L. Altemose, *The Anti-Concurrent Clause and its Impact on Texas Residents after Hurricane Ike*, 16 TEX. WESLEYAN L. REV. 201 (2010).

⁴³ *Id.* (citing *Travelers Indem. Co. v. McKillip*, 469 S.W.2d 160, 162 (Tex. 1971)).

⁴⁴ Erik S. Knutsen, *Confusion About Causation in Insurance: Solutions for Catastrophic Losses*, 61 ALA. L. REV. 957, 977-78 (2010).

⁴⁵ See Randall L. Smith & Fred A. Simpson, *Causation in Insurance Law*, 48 S. TEX. L. REV. 305, 322 (2006).

⁴⁶ *Id.*

incredibly complex method for analysis. Rare is the case when it is clear the precise percent of the loss attributable to a particular peril. For this reason, and others, courts are reluctant to adopt this approach.

III. DEVELOPMENT OF THE APPROACHES IN THE UNITED STATES

In many ways, California is the grandfather of concurrent causation jurisprudence.⁴⁷ A trilogy of California Supreme Court cases has spawned and inspired the jurisprudence throughout the country. *Sabella v. Wisler*, *State Farm v. Partridge*, and *Garvey v. State Farm Fire & Casualty* form the California trilogy. While California would certainly like to disclaim paternity status, the fact remains that other jurisdictions have followed California's lead on many concurrent causation issues.⁴⁸

A. *SABELLA V. WISLER* (EFFICIENT PROXIMATE CAUSE/ DOMINANT CAUSE APPROACH)

In much the same way that California is the grandfather of concurrent causation analysis, *Sabella v. Wisler* is the grandfather of first-party property coverage analysis.⁴⁹ In *Sabella*, a home was damaged by extensive settling, and the settling was caused by a leak in a sewer pipe.⁵⁰ The leaking pipe saturated the fill material surrounding the foundation.⁵¹ The leak was caused by contractor negligence, and more specifically, caused by the contractor inadequately compacting fill material and improperly sealing the sewer pipe joints.⁵² Under the policy, settling was excluded but contractor negligence was covered, and the court was faced with the classic "concurrent causation" question.⁵³ The court reviewed the

⁴⁷ California has certainly tried—and justifiably so—to distance itself as the genesis of concurrent causation. See Joseph Lavitt, *The Doctrine of Efficient Proximate Cause, the Katrina Disaster, Prosser's Folly, and the Third Restatement of Torts: Cracking the Conundrum*, 54 LOY. L. REV. 1, 7 (2008) (citing *Julian v. Hartford Underwriters Ins. Co.*, 110 P.3d 903, 906 (Cal. 2005)). See also Mark D. Wuerfel & Mark Koop, *Efficient Proximate Causation in the Context of Property Insurance Claims*, 65 DEF. COUNS. J. 400, 401 (1998).

⁴⁸ Lavitt, *supra* note 47, at 7.

⁴⁹ 377 P.2d 889 (Cal. 1963).

⁵⁰ *Id.* at 892.

⁵¹ *Id.*

⁵² *Id.*

⁵³ *Id.* at 890.

causes of loss and held that the leaking pipe was the efficient proximate cause of the loss because it set the other events in motion.⁵⁴ The court reasoned that because the efficient proximate cause of the loss was a covered peril, the entire loss was covered.

B. *STATE FARM MUTUAL INSURANCE V. PARTRIDGE* (PRO-POLICYHOLDER/CONCURRENT CAUSATION APPROACH)

The second case in the trilogy is not a property-insurance case, but a liability case.⁵⁵ While it is plainly not a property case, it is included in this discussion because of the confusion generated by the case.⁵⁶

In *Partridge*, the insured was covered by separate automobile and homeowner's insurance policies.⁵⁷ The homeowner's policy provided a much larger coverage amount but excluded losses "arising out of the use" of an automobile.⁵⁸ The facts in *Partridge* were unique: the insured had filed a hair-trigger on a rifle allowing the rifle to be discharged at the slightest touch of the trigger.⁵⁹ The insured and some friends were off-roading hunting jackrabbits when the insured hit a bump; causing the hair-trigger rifle to fire.⁶⁰ The shot hit one of the passengers and caused significant injuries.⁶¹ The trial court found that the insured had committed two negligent acts: the negligent act of filing the hair trigger and the negligent act of driving off-road.⁶² The homeowner's policy covered

⁵⁴ It is important to note the proximate causation issue here and how tort and insurance proximate causation apply. Under tort theories, the loss was proximately caused by the contractor's negligence. This differs under the scope of insurance law, where the goal is to determine the proximate cause of the loss, rather than which culpable party proximately caused the injury.

⁵⁵ *State Farm Mut. Auto. Ins. Co. v. Partridge*, 514 P.2d 123 (Cal. 1973).

⁵⁶ The confusion stems from the misapplication of third-party insurance principles to first-party claims. This is not the only time this issue has confounded courts. See, e.g., Ernest Martin, Jr. & Britton D. Douglas, *The Montrose Case—A Model Loss in Progress Rule Analysis*, available at http://165.97.89.22/files/Uploads/Documents/Attorney%20Publications/Montrose_Case_Progress_Rule_Analysis.pdf (describing the misapplication of the loss in progress rule to first-party losses).

⁵⁷ *Id.*

⁵⁸ *Id.* at 125.

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² Martin & Douglas, *supra* note 56, at 127.

general negligence but excluded damages arising out of the use of the automobile.⁶³

Faced with these multiple perils, the California Supreme Court was again faced with a classic concurrent causation question. The court elected not to follow the precedent in *Sabella* because “the ‘efficient cause’ language is not very helpful, for here both causes were independent of each other: the filing of the trigger did not ‘cause’ the careless driving, nor vice versa.”⁶⁴ Recognizing *Sabella*’s inapplicability to the question at issue, the court developed a new standard for liability losses independent of an analysis of efficient causation. The court held that the fact that “coverage under a liability insurance policy is equally available to an insured *whenever an insured risk constitutes simply a concurrent proximate cause of the injuries.*”⁶⁵ Thus, under *Partridge*, so long as a covered peril substantially contributed to the loss, coverage would be afforded.

While the plain language of *Partridge* clearly limits the case to third-party liability claims, courts began to extend the “concurrent causation” approach to property insurance losses.⁶⁶ For instance, in *Safeco v. Guyton*, the Ninth Circuit analyzed concurrent causation questions after Hurricane Kathleen using the pro-policyholder approach.⁶⁷ The court found that there were two concurrent causes of loss: (a) third-party negligence (a covered loss) in maintaining flood control plans and (b) flood loss (an excluded loss).⁶⁸ The court held that because third-party negligence contributed to the loss, the entire loss was covered, even though the loss was unequivocally caused by flood.⁶⁹

C. *GARVEY V. STATE FARM FIRE & CASUALTY*

Sixteen years, and a mountain of confusion later, the California

⁶³ *Id.* at 126.

⁶⁴ *Id.* at 130 n.10.

⁶⁵ *Partridge*, 514 P.2d at 130 (emphasis added).

⁶⁶ Douglas G. Houser & Christopher H. Kent, *Concurrent Causation in First-Party Insurance Claims: Consumers Cannot Afford Concurrent Causation*, 21 TORT & INS. L.J. 573, 573 (1986).

⁶⁷ *Safeco Ins. Co. of Am. v. Guyton*, 692 F.2d 551 (9th Cir. 1982).

⁶⁸ *Id.* at 554. For a reincarnation of the *Guyton* case post-Hurricane Katrina, see *In re Katrina Canal Breaches Consol. Litig.*, No. 05-4182, 2007 WL 496856, at *2 (E.D. La. Feb. 12, 2007) (holding that losses post-Katrina were not flood losses within the meaning of the water exclusion but, rather, losses resulting from negligence and the breach of the levies).

⁶⁹ These cases were compiled in Houser & Kent, *supra* note 66, at 577-78.

Supreme Court eventually revisited the pro-policyholder concurrent causation approach developed in *Partridge* and rebuked the lower courts for misapplying *Partridge* to property insurance coverage litigation.⁷⁰

Garvey v. State Farm Fire & Casualty involved facts eerily similar to *Sabella*. In the late 1970s the Garveys noticed that an addition to their house was beginning to separate from the main property.⁷¹ The Garveys alleged that the contractor's negligence was the proximate cause of the loss and the loss should be covered.⁷² State Farm responded that settling was the efficient proximate cause of the loss and that any negligence by the contractor was negligible and should not affect coverage.⁷³ Even though the facts were entirely analogous to *Sabella*, the trial court relied on *Partridge*—rather than *Sabella*—and held that the contractor's negligence was a contributing cause, but settling was the dominant cause.⁷⁴ The court held even though negligence was a minor cause and not the efficient proximate cause, the policy should cover the loss because the policy covered negligence.⁷⁵

The California Supreme Court rejected the trial court's application of *Partridge* to property insurance questions.⁷⁶ In first-party property losses, a loss is not necessarily covered just because a covered peril contributes to the loss. Rather, first-party insurance coverage questions require the reviewing court to look at the facts of the case and determine which among the various contributing perils is the "efficient proximate cause" of the loss. The efficient proximate cause has been referred to as "the cause to which the loss is to be attributed, though the other causes may follow it, and operate more immediately in producing the disaster."⁷⁷ Under the efficient proximate cause analysis, if the predominant factor in the loss is covered, the loss is covered even if excluded perils also contribute to the loss. Similarly, if the predominant factor in the loss is excluded, the loss is excluded even if covered perils contribute to the loss.

⁷⁰ *Garvey v. State Farm Fire & Cas. Co.*, 770 P.2d 704 (Cal. 1989).

⁷¹ *Id.* at 705.

⁷² *Id.* at 706.

⁷³ *Id.* at 705-06.

⁷⁴ *Id.* at 704.

⁷⁵ *Id.*

⁷⁶ *Garvey*, 70 P.2d at 713.

⁷⁷ *Id.* at 707 (quoting *Sabella v. Wisler*, 377 P.2d 889, 895 (Cal. 1963) (quoting 6 COUCH ON INS. § 1463 (1930))).

D. SPAWN OF THE TRILOGY AND RAMIFICATIONS

Before *Garvey* could clarify concurrent causation principles, the insurance industry became fearful of the impending onslaught of claims that could be brought as a result of the *Partridge* decision. As discussed above, under some jurisdictions' reading of the policies, if 99.9% of the loss was excluded but 0.01% of the loss was covered, the entire loss could be covered.⁷⁸ Indeed, the insurance industry was justifiably concerned as a broad reading of *Partridge* would prevent insurers from ever excluding certain perils.

As a result of the insurance industry's fears, the industry modified its standard policies in the mid-1980s.⁷⁹ The industry offered revamped standard commercial general liability and commercial property policies, at least in part, to address the concurrent causation decisions spawned by courts applying *Partridge* to property policies.⁸⁰ In order to avoid future *Partridge*-like decisions, the insurance industry included a new exclusion in its standard form contracts.⁸¹ Policies began excluding "loss or damage caused directly or indirectly by any of the following [exclusions]. Such loss or damage is excluded regardless of any other cause or event that contributes concurrently or in any sequence to the loss."⁸² This language has been referred to generally as the "anti-concurrent" policy exclusion.⁸³

IV. COURTS INTERPRETATION OF ANTI-CONCURRENT POLICY EXCLUSIONS

Even though *Garvey* presumably would have corrected the *Partridge*-progeny problems, the anti-concurrent causation exclusions have been interpreted broader than even the insurance industry could have initially imagined.⁸⁴ The development of anti-concurrent causation

⁷⁸ See *supra* Part III.C.

⁷⁹ See Bragg, *supra* note 8, at 392.

⁸⁰ *Id.* at 394.

⁸¹ See *id.*

⁸² See, e.g., Colo. Intergovernmental Risk Sharing Agency v. Northfield Ins. Co., 207 P.3d 839, 841 (Colo. App. 2008) (citing standard ISO policy language).

⁸³ Again, there is a definitional problem here. The clauses are colloquially referred to as anti-concurrent causation clauses, yet they refer to both simultaneous and subsequent causes of loss. See *supra* Part II for further discussion.

⁸⁴ Mark M. Bell, Christopher S. Dunn & James C. Costner, *Confronting Conventional Wisdom on Builders Risk: From Named Insured Status to Concurrent Causation*, 31-Fall CONSTRUCTION LAW 15, 20-21 (2011).

exclusions has proved an especially powerful device for claims denials.⁸⁵ In dealing with anti-concurrent causation exclusions, courts have typically followed one of three approaches: (1) the “freedom of contract” approach, (2) the substantial factor approach, or (3) the Rossmiller/Blue-Pencil approach.⁸⁶

A. FREEDOM OF CONTRACT APPROACH

The freedom of contract approach is probably the most prevalent of the approaches to anti-concurrent causation clauses.⁸⁷ Although many courts have followed this approach, one of the earliest adopters, and one of the clearest analyses on point is found in *Alf v. State Farm Fire and Cas. Co.*⁸⁸

Alf presented the classic chain-of-events concurrent causation question.⁸⁹ The parties agreed that the loss was caused when a pipe on the Alfs’ property ruptured due to unusually low temperatures.⁹⁰ Water then escaped from the ruptured pipe and caused extensive flooding and soil erosion.⁹¹

If *Alf* were decided prior to the 1980s-insurance policy revisions, the policy would have clearly provided coverage. Utah follows the dominant approach to concurrent causation issues, which seeks to find the efficient proximate cause of the loss.⁹² Here, the parties agreed that the efficient proximate cause of the loss was the ruptured pipe—just as in the case of *Sabella*.⁹³ Under the Dominant Cause analysis, which seeks to determine the cause that set the others in motion, the policy would clearly provide coverage.

⁸⁵ Cases expanding on *Partridge* and applying *Partridge* to property insurance questions are colloquially referred to herein as post-*Partridge* decisions.

⁸⁶ No court has actually referred to its approaches by any of these names. But, for clarification and categorization purposes, these names accurately reflect the various approaches taken by U.S. jurisdictions.

⁸⁷ See Michael C. Phillips & Lisa L. Coplen, *Concurrent Causation versus Efficient Proximate Cause in First-Party Property Insurance Coverage Analysis*, 36 BRIEF 32, 35 (2007) (compiling cases throughout the United States).

⁸⁸ 850 P.2d 1272 (Utah 1993); see also *Kane v. Royal Ins. Co. of Am.*, 768 P.2d 678 (Colo. 1989).

⁸⁹ *Alf*, 850 P.2d at 1272.

⁹⁰ *Id.* at 1273.

⁹¹ *Id.*

⁹² See *supra* Part II.

⁹³ See *supra* Part III.A.

Alf, however, was not decided under the pre-1980s insurance policy revisions, and the Utah Supreme Court was faced with the question of whether an insurer could contractually avoid the efficient proximate cause rule. The court held that the efficient proximate cause rule is not an immutable rule of insurance in Utah, but rather, operates as a default rule “only when the parties have not chosen freely to contract out of it.”⁹⁴

The court held that the parties had chosen to contract around the efficient proximate cause rule and that the parties were entitled to do so.⁹⁵ The court reasoned that the anti-concurrent causation language in the policy did not upset norms of reasonable expectations of insureds, and therefore, the contractual modification was permissible.⁹⁶

It is interesting to note that under this interpretation, insurers have essentially turned the tables of *Partridge* on insureds. Courts used *Partridge* as a shield to protect policyholders by granting coverage when there was an argument that the policy should cover the loss.⁹⁷ The freedom of contract approach, conversely, denies coverage when there is an argument that the policy should not cover the loss. This expansion goes far above and beyond the intent of the insurers when they instituted anti-concurrent exclusions.⁹⁸

Today, if the insurer can point to some event in the chain of events that was excluded, the insurer can deny coverage in freedom-of-contract-approach jurisdictions like Utah.⁹⁹ As currently applied, if the insured could argue that 99% of the loss was caused by covered losses, but 1% of the losses was excluded, then the entire loss will be excluded.¹⁰⁰ Additionally, the possibility exists that insurers can modify policies and begin excluding Negligent Acts and Decisions, as an example, in all-risk policies, and thereby effectively prevent coverage for all losses where the loss can be at least partially attributed to someone’s negligence.

In addition to exceeding the insurance industry’s original intent in authoring the exclusions, the expansive scope of the anti-concurrent exclusions is also problematic when considering the nature of the insurance industry. Most insurance policies are contracts of adhesion incapable of

⁹⁴ *Alf*, 850 P.2d at 1277.

⁹⁵ *Id.* at 1272.

⁹⁶ *Id.* at 1278.

⁹⁷ Houser & Kent, *supra* note 66, at 575-77.

⁹⁸ See *supra* note 79-83 and accompanying text.

⁹⁹ *Davidson Hotel Co. v. St. Paul Fire and Marine Ins. Co.*, 136 F. Supp. 2d 901 (W.D. Tenn. 2001).

¹⁰⁰ See *Chattanooga Bank Assoc. v. Fid. & Deposit Co.*, 301 F. Supp. 2d 774 (E.D. Tenn. 2004).

modification by individual policyholders. These disperse policyholders do not possess the lobbying powers or the contractual capacity or influence to cause ubiquitous changes across all policy lines of insurance, which is why the policies remain as they are today. The insurance industry can effectively modify the policy in response to negative precedent; whereas the diverse policyholders do not possess similar power. Accordingly it should be incumbent on either the courts or legislature to prevent over-expansive use of the anti-concurrent exclusions—especially when the interpretations exceed the intended purpose of the exclusions.

B. SUBSTANTIAL FACTOR APPROACH

Recognizing the various problems associated with the freedom-of-contract approach, some courts have held that in order for anti-concurrent exclusions to apply, the excluded loss must be a substantial factor or the efficient proximate cause of the loss.¹⁰¹ There are four states that have followed this approach and expressly rejected the freedom of contract approach. California¹⁰² and North Dakota¹⁰³ have done so by code and Washington and West Virginia have done so by case law.¹⁰⁴

The first case to reject the freedom-of-contract approach without relying on insurance code regulations was *Safeco Insurance Co. v. Hirschmann*.¹⁰⁵ In *Hirschmann*, severe winds were followed by heavy rains

¹⁰¹ Phillips & Coplen, *supra* note 87, at 35 (compiling cases throughout the United States).

¹⁰² CAL. INS. CODE §§ 530, 532 (West 2005) (“An insurer is liable for a loss of which a peril insured against was the proximate cause, although a peril not contemplated by the contract may have been a remote cause of the loss; but he is not liable for a loss of which the peril insured against was only a remote cause.”).

¹⁰³ N.D. CENT. CODE § 26.1-32-01 (2010) (“An insurer is liable for a loss proximately caused by a peril insured against even though a peril not contemplated by the insurance contract may have been a remote cause of the loss. An insurer is not liable for a loss of which the peril insured against was only a remote cause. The efficient proximate cause doctrine applies only if separate, distinct, and totally unrelated causes contribute to the loss.”).

¹⁰⁴ At the time of this writing, the Colorado Supreme Court has granted certiorari to determine whether a loss is covered where 90% of the loss was covered and 10% was excluded. *Colo. Intergovernmental Risk Sharing Agency v. Northfield Ins. Co.*, 207 P.3d 839 (Colo. App. 2008), *cert. granted*, No. 08SC907, 2009 WL 1485804 (Colo. May 26, 2009).

¹⁰⁵ 773 P.2d 413 (Wash. 1989). For a recent example, see *Sprague v. Safeco Ins. Co. of Am.*, 241 P.3d 1276, 1278 (Wash. Ct. App. 2010) (“In analyzing coverage, Washington follows the efficient proximate cause rule. Under this rule,

and landslides.¹⁰⁶ The Hirschmanns' home was pushed from its foundation and completely destroyed because of strong winds and water saturation of the soil.¹⁰⁷ According to one expert, the primary cause of the hillside's collapse was the heavy rainfall.¹⁰⁸

Safeco denied coverage and conceded during the proceedings that if the policy had been interpreted prior to the 1980 policy revisions, then the loss would have been covered in Washington.¹⁰⁹ Safeco argued that even though Washington adheres to the dominant-cause approach, the post-1980s policy revisions overcome the dominant-cause approach.¹¹⁰ Safeco argued that it should be able to exclude coverage since at least part of the loss was excluded.¹¹¹

The court in *Hirschmann* rejected Safeco's argument and held that the efficient proximate cause rule represents an immutable principle of Washington insurance law, and that the parties cannot contract around it.¹¹² The court held that because the primary causes of the loss included the covered perils of wind and rain, the entire loss was covered.¹¹³

Murray v. State Farm Fire & Cas. Co. provides a thorough primer on concurrent causation including a description of the resulting disproportionate forfeiture if the efficient proximate cause rule is ignored.¹¹⁴ In *Murray*, State Farm argued that its anti-concurrent clause "operates to defeat the efficient proximate cause doctrine."¹¹⁵ Further, State Farm "argue[d] that if earth movement *in any way contribute[d]* to a loss, regardless of the proximate cause, then under the lead-in [anti-concurrent] clause the entire loss is excluded from coverage."¹¹⁶

The court in *Murray* rejected State Farm's contention and captured the essence of potential problems associated with abandoning the efficient

the predominant cause of the loss determines coverage.") (footnotes omitted); *but cf.* *City of Everett v. Am. Empire Surplus Lines Ins. Co.*, 823 P.2d 1112, 1115 (Wash. Ct. App. 1991) (holding that insurance policies that use the phrase "arising out of" do not warrant efficient proximate cause analysis).

¹⁰⁶ *Hirschmann*, 773 P.2d at 413.

¹⁰⁷ *Id.* at 414.

¹⁰⁸ *Id.*

¹⁰⁹ *Id.*

¹¹⁰ *Id.*

¹¹¹ *Id.* at 413-14.

¹¹² *See Hirschmann*, 773 P.2d at 415-16.

¹¹³ *Id.* at 417-18.

¹¹⁴ 509 S.E.2d 1 (W. Va. 1998).

¹¹⁵ *Id.* at 14.

¹¹⁶ *Id.* (emphasis added).

proximate cause rule in the face of anti-concurrent causation clauses:

Indeed if we were to give full effect to the State Farm policy language excluding coverage whenever an excluded peril is a contributing or aggravating factor in the loss, we would be giving insurance companies *carte blanche* to deny coverage in nearly all cases.¹¹⁷

Applying these principles, the West Virginia court rejected a broad reading of anti-concurrent clauses and held that the efficient proximate cause rule cannot be modified to abandon the reasonable expectations of the insured.¹¹⁸ The court held that the reasonable insurer expects to have losses covered where the predominant cause of the loss is covered.¹¹⁹

One could certainly criticize the substantial factor approach because it essentially ignores the 1980s revisions to insurance policies and renders the anti-concurrent policy ineffective. As demonstrated by the Washington and West Virginia cases, these courts essentially apply the same analysis that they applied before the introduction of anti-concurrent causation clauses. Opponents to the approach, including insurers in general, argue that courts applying the substantial factor approach make anti-concurrent clauses superfluous and meaningless.

In this author's opinion, the criticism is unproblematic. Insurers introduced the anti-concurrent causation clauses to combat post-*Partridge* expansion of concurrent causation. Michael E. Bragg, assistant counsel for State Farm Insurance, wrote an article in the 1980s that discussed State Farm's specific attempts to draft policy language to avoid post-*Partridge* concurrent causation interpretations.¹²⁰

The difficulty of the industry's task in combating concurrent causation embraces two distinct but related issues intertwined in the court decisions. First, the courts are creating new "causes" of loss never contemplated by property insurance policy drafters. Most important of these causes are negligence and other human conduct. Such

¹¹⁷ *Id.* (emphasis added) (quoting *Howell v. State Farm Fire & Cas. Co.*, 218 Cal. App. 3d 1446, 1456 n.6 (Cal. Ct. App. 1990)).

¹¹⁸ *Id.* at 14-15.

¹¹⁹ *Id.*

¹²⁰ Michael E. Bragg, *Concurrent Causation and the Art of Policy Drafting: New Perils for Property Insurers*, 20 FORUM 385 (1985).

conduct may be active, passive, willful, negligent, imprudent, untimely, or any other word which describes how people act or fail to act. Second, the courts are telling us that the proper causation standard is no longer to attribute the loss to a single proximate cause, but rather to grant coverage if *any* of the causes of the loss has not been specifically excluded.¹²¹

As demonstrated by this influential article, anti-concurrent causation clauses were not intended to impact the efficient proximate causation standard employed by post-*Sabella* interpretations. The *Sabella* analysis seems to offer a fair and reasonable interpretation from both the insurer and policyholder perspectives. Insurers intended to prevent post-*Partridge* interpretations where the loss was covered if the insured could point to a single factor that contributed to the loss. While the case has not yet arisen in any jurisdictions following the pro-policyholder approach, the case can be made that the anti-concurrent exclusions would be effective in those jurisdictions and would move those jurisdictions from a post-*Partridge* analysis to a post-*Sabella* analysis.

Accordingly, by applying the substantial factor approach, as Washington and West Virginia courts have, insurers are adequately safeguarded against post-*Partridge* interpretations. Additionally, the approach mitigates the potential for insurers to deny losses when the loss was proximately caused by a covered cause.

C. THE ROSSMILLER/BLUE PENCIL APPROACH

The third approach that courts have used employs a much more involved and detailed analysis of concurrent causation. It seems that *Corban v. USAA* is the only court to have used this approach to date, but I have included it as its own approach, because it is quite likely another court will follow Mississippi's lead. This approach has largely evolved from the work of concurrent causation scholar/practitioner David Rossmiller.¹²²

¹²¹ *Id.* at 389.

¹²² David P. Rossmiller, *Katrina in the Fifth Dimension: Hurricane Katrina Cases in the Fifth Circuit Court of Appeals*, in NEW APPLEMAN ON INSURANCE: CURRENT CRITICAL ISSUES IN INSURANCE LAW 71, 86 (Matthew Bender ed., 2008) [hereinafter "Rossmiller, *Katrina*"]; David P. Rossmiller, *Interpretation and Enforcement of Anti-Concurrent Policy Language in Hurricane Katrina Cases and Beyond*, in NEW APPLEMAN ON INSURANCE: CURRENT CRITICAL ISSUES IN

Rossmiller published two influential articles in 2007 and 2008, which cogently argue for a particular interpretation of “concurrent” when used in relation to concurrent causation.¹²³ Under Rossmiller’s view, concurrent should either refer to perils (a) acting in coordination or (b) acting in sequence.¹²⁴ For instance, assume that a fire and earthquake both operated to cause a loss: (a) acting in coordination would occur if the earthquake worked in conjunction with the fire to cause the same damage; (b) acting in sequence would occur if the fire resulted from the earthquake; and (c) a non-concurrent result would occur if the fire merely occurred at the same time as the earthquake but was not brought about by the earthquake.¹²⁵

According to Rossmiller, Hurricane Katrina did not actually involve concurrent causes of loss “not because they came at different times, but because each force acted separately to create unique damage”¹²⁶ – as in the third earthquake/fire example described above. Under Rossmiller’s view, the fact that both wind and flood were “products of the same larger phenomenon, a hurricane, is irrelevant.”¹²⁷ The argument follows that losses are concurrent only where multiple causes produce the same damage, and losses are not concurrent when multiple causes result in multiple losses.

While Rossmiller’s articles have been cited by other courts,¹²⁸ the first court to adopt his approach was the Mississippi Supreme Court. The Mississippi Supreme Court addressed the concurrent causation question for the first time in *Corban v. USAA*¹²⁹ after several federal courts had provided *Erie*-guesses as to how Mississippi would analyze concurrent

INSURANCE LAW 43, 65 (Matthew Bender ed., 2007) [hereinafter “Rossmiller, *Interpretation*”].

¹²³ See generally Rossmiller, *Katrina*, *supra* note 122; Rossmiller, *Interpretation*, *supra* note 122.

¹²⁴ See generally Rossmiller, *Katrina*, *supra* note 122; Rossmiller, *Interpretation*, *supra* note 122.

¹²⁵ The earthquake/fire analogy is used throughout this article. For references to insurance/earthquakes, the reader should ignore any potential differing results that would occur under an analysis of the New York Standard Fire Policy. For purpose of the analogy, assume that neither New York’s nor any other state’s standard fire policies apply.

¹²⁶ Rossmiller, *Interpretation*, *supra* note 122, at 65.

¹²⁷ *Id.*

¹²⁸ Colo. Intergovernmental Risk Sharing Agency v. Northfield Ins. Co., 207 P.3d 839, 842 (Colo. App. 2008).

¹²⁹ Corban v. United Servs. Auto Ass’n, 20 So. 3d 601 (Miss. 2009).

causation questions.¹³⁰

The Corbans owned a two-story home that was damaged—but not destroyed—by Hurricane Katrina.¹³¹ USAA inspected the home and determined that although the wind caused some damage to the roof and second floor, the majority of damage to the first floor was caused by flooding.¹³² Accordingly, USAA paid the portion of damages to the roof and second floor related to the wind damage and denied coverage for the first floor because of the anti-concurrent flood exclusion.¹³³

In order to determine whether the denial was proper, the court in *Corban* narrowly defined concurrent.¹³⁴ Although there are numerous definitions that the court could have used to define concurrent, *Corban* used the following narrow definition: the “exclusion applies only in the event that the perils [1] act in conjunction, [2] as an indivisible force, [3] occurring at the same time, [4] to cause direct physical damage resulting in loss.”¹³⁵

Additionally, the court held that the provision “in any sequence” irreconcilably conflicts with Mississippi law and is void and unenforceable.¹³⁶ By rejecting the “in any sequence language” in the anti-concurrent exclusion, the court also addressed questions brought up by federal courts and held that “[a]n insurer cannot avoid its obligation to indemnify the insured based upon an event which occurs subsequent to the covered loss.”¹³⁷

Under the narrow definition of concurrent, the insurer has the

¹³⁰ For a fascinating history of the chronology of the federal courts *Erie*-guess analogies, see Rossmiller, *Katrina*, *supra* note 122; see also Bell et al., *supra* note 84, at 21-23.

¹³¹ See *Corban*, 20 So. 3d at 605-06.

¹³² *Id.* at 606.

¹³³ *Id.*

¹³⁴ *Id.* at 614.

¹³⁵ *Id.* (brackets added for clarity).

¹³⁶ *Id.* at 615. The court presumably could have declared the entire exclusion void as a result of this provision, but for reasons unexplained by the court, the court seems to have severed this provision from the rest of the exclusion.

¹³⁷ *Corban*, 20 So. 3d at 613. Indeed, USAA also rejected the Fifth Circuit’s analogy when pressed during trial. According to USAA, “if an insured’s roof is breached and rainwater comes in, damaging a carpet, USAA pays for rainwater damage to the carpet . . . even if storm surge subsequently. . . destroy[s] the carpet.” *Id.* at 610; for additional discussion on indemnification and subrogation, see Jay S. Bybee, *Profits in Subrogation: An Insurer’s Claim to be More than Indemnified*, 1979 BYU L. REV. 145 (1979).

burden of proving that two perils operate in conjunction and that the perils operated contemporaneously. In *Corban*, and likely the majority of Katrina claims, the wind and the flood did not operate contemporaneously or in conjunction because most experts estimate that the wind preceded the flooding by up to four hours.

Corban also established the relevant burdens of proof for insurance claims. Under an all risk policy, the insured has the burden to prove that a loss occurred. After proving that a loss occurred, the burden shifts to the insurer to prove an affirmative defense—for example, demonstrating that the peril is excluded under the policy. In *Corban*, it was clear that a loss occurred; therefore USAA had the burden of proving by a preponderance of evidence that the damages were caused by the excluded peril of flooding.¹³⁸

I refer to this approach as the “blue pencil” approach because it strikes a portion of the anti-concurrent exclusion, but does not invalidate the entire clause. As stated previously, anti-concurrent exclusions generally exclude losses caused concurrently and “in any sequence.” *Corban* held that the “in any sequence” language was unenforceable, but held that anti-concurrent clauses are enforceable. While this represents a more policyholder-friendly approach than the courts that simply enforce anti-concurrent causation clauses wholesale, it still leaves open the possibility that anti-concurrent causation exclusions can exclude losses where 99% of the loss is covered but 1% of the loss is excluded.

V. A CALL FOR CLARITY AND A REVISION OF THE TERMS OF INTERPRETATION

As stated at the outset, and as evidenced by the approaches to concurrent causation and anti-concurrent causation exclusions, the nomenclature of concurrent causation has become so bastardized that the concurrent and efficient proximate cause issues have become an untraceable mess.

To correct this mess, courts and commentators should re-visit concurrent causation to redefine the terms to more accurately reflect the underlying policies and provide additional clarity. In addition to redefining the relevant concurrent causation terms, courts, insurers, and policyholders, should take a new approach to analyzing concurrent causation questions.

¹³⁸ *Corban*, 20 So. 3d at 618-19.

A. DEFINITIONAL CLARITY

There are two terms that proliferate “concurrent causation” analyses and courts continuously apply these definitions inappropriately: “concurrent causation” and “efficient proximate cause.”

As currently defined, these terms are inexact and create confusion and result in inconsistent application. “Concurrent” is used (1) to refer to any multi-factor causation analysis,¹³⁹ (2) to refer to a particular type of multi-factor causation analysis,¹⁴⁰ and (3) as a method or approach to multi-cause losses.¹⁴¹ Similarly, “efficient proximate cause” is used (1) as a method or approach to multi-cause losses,¹⁴² (2) as the “moving cause of loss” when there is a chain-of-events preceding a loss,¹⁴³ and (3) as the “predominant” factor in non-chain-of-event losses when multiple perils combine to cause a loss.¹⁴⁴

Given the conflation of terms, it is time to redefine these terms to allow greater accuracy and precision. Additionally, given the current confusion generated by the term “concurrent”, courts, commentators, and insurers should drop the term “concurrent” from the insurance lexicon.¹⁴⁵

In order to provide clarity on “concurrent causation” questions, the term concurrent causation must be addressed first. Although by definition concurrent requires temporal proximity,¹⁴⁶ the term has been eviscerated to the point that concurrent no longer has any definitional meaning. To demonstrate this point, Rossmiller, one of the most well-versed and persuasive writers on the subject, argues that temporal proximity—the essence of concurrence—is “irrelevant” to the question of whether a loss is concurrent.¹⁴⁷ If “concurrent” does not relate to temporal proximity, then no concurrent causation analysis can truly be said to be necessarily related to concurrence. Correspondingly, when courts attempt to define the term,

¹³⁹ See, e.g., Bragg, *supra* note 120, at 285.

¹⁴⁰ See, e.g., Garvey v. State Farm Fire & Cas. Co., 770 P.2d 704, 709 (Cal. 1989).

¹⁴¹ See *supra* Part II.

¹⁴² See, e.g., Lavitt, *supra* note 47, at 2.

¹⁴³ See, e.g., Sabella v. Wisler, 377 P.2d 889, 895 (Cal. 1963).

¹⁴⁴ See, e.g., Shinrone, Inc. v. Ins. Co. of N. Am., 570 F.2d 715, 718-19 (8th Cir. 1978).

¹⁴⁵ Perhaps the word need not be dropped permanently, but certainly a long hiatus would be beneficial to avoid the current conflation of terms currently applied to “concurrent.”

¹⁴⁶ See sources cited *supra* note 3.

¹⁴⁷ See *supra* note 122 and accompanying text.

they are unable to appropriately define concurrent while maintaining some semblance of the term as defined by dictionaries or as originally intended by courts and insurers.

For that reason, courts and commentators should avoid using the term “concurrent” to refer to a loss caused by more than one cause. Instead, courts and commentators should use “multi-cause” in its place. Either a loss is caused by one cause, or the loss is caused by multi-causes. If the loss is caused by one cause, the analysis is simple and the court determines whether that loss is covered. If, however, the loss is caused by multi-causes, then courts should engage a new approach to the multi-cause loss analysis.

This presents a simple remedy to an unnecessarily complicated problem. There is no reason that multi-cause losses should be referred to as concurrent, but that is what has been done for years. If there were some reason to use the term concurrent, I would refrain from suggesting a replacement. However, there is absolutely no reason whatsoever to refer to a loss caused by multiple causes as a “concurrent” loss.

Second, the term efficient proximate cause has been used in so many different ways that there is confusion about its definition as well. As originally envisioned, efficient proximate cause related to chain-of-event questions.¹⁴⁸ Originally, courts would attempt to determine the efficient proximate cause to decide which event set the other events in motion.¹⁴⁹ Thus, the precise definition for efficient proximate cause is the cause that sets the others in motion and relates expressly to chain-of-event losses. This is the only place where the term efficient proximate cause should be used.

Over time, courts and commentators began to use efficient proximate cause more loosely and applied the term to non-chain-of-event multi-cause losses. Efficient proximate started being defined as the “predominant factor” in a loss and has been used to refer to the dominant-cause approach.¹⁵⁰ This has generated confusion because courts now attempt to look for the “moving” cause of loss even when there is not a chain-of-events preceding the loss. For non-chain-of-event losses, however, there is no “moving” cause of loss and courts must look to the predominant or substantial cause of the loss.

¹⁴⁸ Lavitt, *supra* note 47.

¹⁴⁹ Vintila v. Drassen, 52 S.W.3d 28, 41-42 (Mo. Ct. App. 2001).

¹⁵⁰ Elizabeth L. Perry, *Why Fear the Fungus? Why Toxic Mold is and is not the Next Big Toxic Tort*, BUFF. L. REV. 257, 280 (2004).

B. NEW ANALYSIS FOR MULTI-CAUSE LOSSES

Rather than having courts attempt to analyze insurance policies in a vacuum, I would propose that courts and commentators address concurrent causation issues using a more methodical, categorical approach.

In insurance coverage, categorization is often essential to understanding the issues. Indeed, the proliferation of confusion concerning multi-cause losses can be traced to deficiencies in categorization. For instance, the post-*Partridge* proliferation largely occurred because courts failed to appropriately categorize the losses. Different concerns arise in property and liability disputes and courts should treat the disputes differently.¹⁵¹ In the post-*Partridge* era, courts failed to properly distinguish property from liability cases and inappropriately applied liability standards to property cases.

To avoid these types of categorical problems, this article advocates a more methodical approach and recommends that courts engage in an analysis using a number of discreet, step-by-step questions. The discreet questions would encourage courts to appropriately categorize the loss and subsequently apply the proper means of analysis to that particular category of loss. This approach would more uniformly address multi-cause losses and would lead to improved consistency and efficiency throughout jurisdictions, would avoid inequitable results, and would lead to greater contract certainty.¹⁵²

When addressing insurance coverage questions, the key concern is causation and whether the peril causing the loss is covered or excluded. The approach advocated in this article presents a more direct-line, causal, approach to causation questions than the piecemeal approach currently employed by the courts.

Obviously the threshold question in a coverage dispute concerns the determination of what specific peril or perils contributed to the loss. When losses only involve one peril, the analysis is straightforward: was the peril covered or excluded? Conversely, when losses involve multiple

¹⁵¹ *Garvey v. State Farm Fire and Cas. Co.*, 770 P.2d 704, 710 (Cal. 1989) (“Liability and corresponding coverage under a third-party insurance policy must be carefully distinguished from the coverage analysis applied in a first-party property contract. Property insurance, unlike liability insurance, is unconcerned with establishing negligence or otherwise assessing tort liability.”) (citing Michael E. Bragg, *Concurrent Causation and the Art of Policy Drafting: New Perils for Property Insurers*, 20 FORUM 385, 386 (1985)).

¹⁵² As with most insurance questions, these issues are often best addressed by a flowchart and I have attached the flowchart to the appendix for review.

causes, the analysis becomes far more complicated. Accordingly, once the stakeholders recognize that the loss involves multiple causes, it would behoove the courts to address a series of questions before opining on the resulting coverage question: (1) did the causes operate in an unbroken chain of events or did the causes operate independently?; (2) if the losses operated independently, did they act simultaneously or sequentially?; (2a) if the losses were simultaneous, were the various causes independently sufficient or independently insufficient to cause the loss?; (2b) if the losses were sequential, what cause and resultant loss came first and did the second cause exacerbate the preceding loss?¹⁵³

1. Did the Causes Operate in an Unbroken Chain or did the Causes Operate Independently?

Different analyses are required when dealing with losses caused by an unbroken chain-of-events and losses caused by independent perils. For unbroken chains-of-events, courts typically try to determine what was the “moving” cause of the loss, or stated in other terms, “if the immediate cause of the loss was dependent on other forces or events, then the trier of fact [is] required to engage in a process of selection to determine the ‘efficient’ cause of the loss.”¹⁵⁴

If there is a chain-of-events, the court should look to the “efficient proximate cause of the loss.” In typical chain-of-event scenarios, the event that sets the others in motion is well established and easily ascertainable. For instance, in a relatively recent Ninth Circuit Court of Appeals case, the parties unequivocally agreed on the efficient proximate cause of the loss stemming from an unbroken chain of events.¹⁵⁵ In *Terminal Freezers*, the policyholder had built a commercial freezer facility.¹⁵⁶ Eventually, the policyholder discovered that ice was accumulating in the ceilings and walls.¹⁵⁷ The parties unanimously agreed that the ice was caused due to an unbroken chain-of-events.¹⁵⁸ During construction, the contractor had

¹⁵³ See flowchart attached to appendix for clarity on the steps.

¹⁵⁴ Lavitt, *supra* note 47, at 9 (citing *Gen. Mut. Ins. Co. v. Sherwood*, 55 U.S. 351, 366-67 (1852)). California has certainly tried—and justifiably so—to distance itself as the genesis of concurrent causation. *Id.* at 7.

¹⁵⁵ See *Terminal Freezers Inc. v. U.S. Fire Ins.*, No. 08-35623, No. 08-35656, 2009 U.S. App. LEXIS 20321 (9th Cir. Sept. 3, 2009).

¹⁵⁶ *Terminal Freezers Inc. v. U.S. Fire Ins.*, No. C07-0090BHS, 2008 U.S. Dist. LEXIS 48280, at *3 (W.D. Wash. June 23, 2008).

¹⁵⁷ *Id.* at *20.

¹⁵⁸ *Id.* at *20-21.

defectively installed a vapor barrier, the defective vapor barrier allowed water vapor to enter the facility, the water vapor infiltrated ceiling tiles and insulation, and the water vapor then froze in the ceiling tiles and insulation; thereby destroying the interior of the facility and causing significant damage.¹⁵⁹ In *Terminal Freezers*, the immediate cause of the loss was water vapor freezing; however, there was no doubt between the parties that the real “cause”, or the “efficient” cause of the loss, was the defectively installed vapor barrier: but-for the defectively installed vapor barrier, the ice would not have accumulated in the building.¹⁶⁰ Like *Terminal Freezers*, most chain-of-event cases provide a relatively straightforward question that is often capable of agreement between the parties.

Accordingly, for chain-of-event cases, the court should continue to seek to determine the efficient proximate cause of the loss and determine whether the efficient proximate cause is covered or excluded. If the cause is covered, the entire loss should be covered; conversely if the efficient proximate cause is excluded, then the entire loss should be excluded.

In non-chain-of-event cases, however, there is no “efficient proximate cause” setting in motion an unbroken chain of events. Accordingly, the efficient proximate cause analysis is inappropriate for independent cause cases, which helps to explain why courts have had such difficulty attempting to fit the efficient proximate cause framework into independent causation analyses. Thus, courts should employ an entirely different analysis when evaluating these types of losses. In these cases, courts should determine whether the causes operated simultaneously or sequentially.

2. Did the Causes in the Multi-Cause Loss operate Simultaneously or Sequentially

Simultaneity is important in the insurance context. Modern “concurrent causation”—multi-cause—jurisprudence arose when California addressed a loss where simultaneous causes operated to create the loss.¹⁶¹

In this author’s view, perils operating simultaneously should be

¹⁵⁹ *Id.* at *4.

¹⁶⁰ *Terminal Freezers*, 2008 U.S. Dist. LEXIS 48280 (the parties did dispute, however, whether the resultant ice formations should be covered or excluded, but the case is illustrative of how courts employ the chain-of-events analysis works).

¹⁶¹ *State Farm Mut. Auto. Ins. Co. v. Partridge*, 514 P.2d 123 (Cal. 1973).

analyzed separately from independent perils operating sequentially.¹⁶² The temporal differences raise independent questions. Just as in the case of conflating chain-of-event and independent losses, when courts and commentators begin classifying simultaneous and sequential losses together, confusion results because the concerns and the issues in both cases are separate and distinct.

a. Simultaneous Losses

For perils operating simultaneously, courts should first determine whether the various perils were independently sufficient or independently insufficient to cause the loss. By way of analogy, the quintessential independent-simultaneous-multi-cause loss would present itself if an earthquake (excluded peril) occurred at the *exact* same time as a fire (covered peril). For this analogy, the two events are entirely unrelated, and the property was completely destroyed as a result of the loss.

In this analogy, the court would determine whether a covered peril was independently sufficient to cause the entire loss. If a covered peril is sufficient to cause the entire loss, then the entire loss should be covered. For example, in this analogy, if the fire could have caused the entire loss, then the loss should be covered, even if the earthquake could also have caused the entire loss.

If the covered peril was not sufficient to cause the entire loss, and the excluded peril could have independently caused the loss, then the entire loss should be excluded. Continuing the analogy, if the fire could not have caused the entire loss, but the earthquake could have caused the entire loss, then the entire loss should be excluded.

If, however, neither the fire nor the earthquake could have independently caused the loss, then the court should determine which of the two perils was the “predominant” cause of loss.¹⁶³ If the court determines that the fire is the predominant cause of loss, then the entire loss should be

¹⁶² It is important to note that once we are in step 2, sequential losses do not refer to sequential unbroken chain-of-events. Rather, sequential refers solely to independent perils occurring sequentially. For chain-of-event losses, courts should continue applying the efficient proximate cause analysis as discussed *supra* Part V.B.1.

¹⁶³ It is important to note the definitional consistency that needs to be employed in this category. This analysis should be referred to as seeking the “predominant” cause of the loss. This should not be referred to as the “efficient proximate cause” of the loss because that term is limited to chain-of-event situations, which are not present in this example.

covered. If however, the court determines that the earthquake is the predominant cause of the loss, then the entire loss should be excluded.

The rationales for this approach relate to reasonableness and notions of fairness. If one covered cause of loss was sufficient to cause the entire loss, then the insurer should not benefit from the fortuitous circumstance that an excluded loss operated at the same time. The insurer underwrites the policy and intends to provide insurance for certain events. Once that event is triggered, the insurer should not be able to benefit because an additional cause occurred at the same time. The policyholder pays a premium for particular coverages, and once those coverages are triggered, the insurer is obligated to pay. Conversely, if an excluded peril could have caused the entire loss, then the policyholder should not be able to benefit when the property would have been completely destroyed and the damages caused by the covered perils were less than the damages caused by the excluded perils. Similarly, if neither peril could have independently caused the loss, fairness dictates that the court should attempt to determine which cause was the predominant cause of the loss. If the predominant factor in the loss was excluded, the policyholder should not be able to receive coverage when the bulk of the damage is caused by excluded causes. By that same token, the insurer should not be able to avoid coverage when covered losses predominate.

b. Sequential Losses

For *independent* causes occurring in sequence, the threshold question should attempt to determine which cause and resultant loss came first. The second question would ask whether the subsequent loss exacerbated the damage or created new damage.

While some courts have ignored the sequence of losses, fundamental notions of insurance dictate that the sequence is essential to determine whether there should be coverage. As prudently stated by the Mississippi Supreme Court:

No reasonable person can seriously dispute that if a loss occurs, caused by either a covered peril (wind) or an excluded peril (water), that particular loss is not changed by any subsequent cause or event. Nor can the loss be excluded after it has been suffered, as the right to be indemnified for a loss caused by a covered peril attaches at that point in time when the insured suffers deprivation of, physical damage to, or destruction of the property insured.

An insurer cannot avoid its obligation to indemnify the insured based upon an event which occurs subsequent to the covered loss.¹⁶⁴

In *Corban*, the court addressed immutable principles of insurance coverage. Covered losses do not become excluded merely because a subsequent cause operates on the loss.¹⁶⁵ Similarly, excluded losses do not become covered merely because subsequent covered perils happen to impact the loss. Thus, the key question should center on which peril came first and whether that peril is covered or excluded: If the peril is covered, the loss should be covered, and the reverse holds true as well.

After determining which loss came first, the court should determine whether the subsequent peril exacerbated the loss or created new damage. If the subsequent cause exacerbated the loss, then the exacerbated damages should be categorized according to the prior loss. If, however, the subsequent cause creates new damage, then the court should re-analyze whether that cause is covered or excluded and provide coverage for the new loss accordingly.

For example, if an earthquake (excluded) damaged a property and two hours later a fire (covered) came and merely exacerbated the earthquake damage, the entire loss would be excluded. If, however, the earthquake damaged the foundation of the property causing distinct damages, and the fire later damaged the roof and framing, then the fire damage should not be excluded merely because an earthquake caused some damage to the property.

The rationale for this approach relates to reasonableness and doctrines of fairness. It should be an immutable doctrine of insurance coverage that covered losses do not become uncovered merely because the insurer has not had yet paid the claim.

By way of analogy, suppose a policyholder suffered a loss on the 1st of the month, and the insurer acknowledged the loss was covered and payment should be made on the policy. No reasonable insurer would argue that the covered loss on the 1st of the month becomes excluded simply because the policyholder suffers a subsequent loss caused by an excluded peril on the 31st of that month. Although that analogy seems absurd, that is essentially the argument that insurers make during sequential multi-cause

¹⁶⁴ *Corban v. United Servs. Auto. Ass'n*, 20 So. 3d 601, 613 (Miss. 2009) (footnotes omitted).

¹⁶⁵ *Id.*

losses.¹⁶⁶ These arguments should be rejected when an excluded peril merely subsequently impacts the property and causes the same damage or merely exacerbates the previous loss. Just as the argument is rejected for losses occurring 30 days later, so too should they be rejected when occurring 30 minutes later.

Similarly, the policyholder should not be able to receive an undue benefit. By analogy, if an automobile is in an accident and the hood is mangled and unfit for daily use, no reasonable policyholder would argue that the policyholder should be able to recover for an unrelated key-scratch on the same hood. The same analogy applies. The fact that a covered event occurs after an excluded event should not morph the excluded loss into a covered one.

The area for potential pushback in this approach concerns the exacerbation/new loss distinction. If the subsequent loss significantly exacerbates the loss, the stakeholders may have a claim that there should be some offset. However, experience indicates that bifurcating losses is extremely difficult, and apportionment is inexact and difficult to prove. The problem only becomes more complicated in cases of total losses. Thus, for clarity and policy consistency, a subsequent exacerbation of a previous loss should not affect the prior loss determination.

In cases where the losses and subsequent causes can be clearly bifurcated, the subsequent loss should be analyzed under general principles of insurance interpretation.

Revisiting the earthquake-fire analogy, if an earthquake were to damage the foundation and then an unrelated fire were to strike the property, the damage from the earthquake would clearly be excluded since it occurred first. If the earthquake caused the total loss of the property, then the loss would be excluded, even if a subsequent unrelated fire struck the location and would have assuredly burned the building to the ground. If the earthquake did not cause a total loss of the property, and an unrelated fire later struck the same property, then the court would look to the impact on the property and the nature of the fire damage. If the fire damage exacerbated structural problems caused by the earthquake, then the resultant fire-structural damage would be excluded. Conversely, if the fire damaged property was undamaged by the earthquake, then the policy would cover the resultant unrelated fire damage.

¹⁶⁶ The ISO policy exclusions exclude losses caused “in any sequence” by excluded perils. Tim Ryles, *Rethinking Concurrent Causation and the Flood Exclusion: Further Comments on Katrina-Related Coverage Disputes*, IRMI.COM, Sept. 2007, <http://www.irmi.com/expert/articles/2007/ryles09.aspx>.

C. ARGUMENTS AGAINST THE REVISED APPROACH

As for my proposal calling for definitional clarity, most courts and commentators would probably agree that the current definitional landmine is unworkable and that it is time to revisit the terms relating to these issues. While some may disagree with the terms used in this article, most commentators will probably agree that the current lexicon is unworkable.

As for the approach this article takes with respect to jurisprudential analysis, there are probably two main areas for attack: (1) the article essentially adopts the dominant cause approach in most circumstances and would result in a pro-policyholder jurisprudential shift and (2) the revised approach could create additional confusion.

1. The Dominant-Cause or Efficient Proximate Cause Critique

In many ways, the approach advocated for in this article does adopt some iterations of the dominant-cause approach. However, this incorporation is intentional: (1) when insurers began inserting anti-concurrent causation clauses into insurance policies, the insurers were trying to combat post-*Partridge* analyses to multi-cause losses; and (2) policyholders do not possess the same negotiating leverage or coordination of effort to institute the reasonable changes proffered in this article.

First, insurers sought to avoid situations where a minor covered cause in a chain-of-events operated to cover the entire loss.¹⁶⁷ By applying the approach advocated in this article, the insurer is back in the pre-*Partridge* analysis of multi-cause losses. In the perfect world, there would be much greater uniformity across jurisdictions, which would allow insurers to be able to maintain some sense of contractual certainty. Insurers would know ex ante how courts would address multi-cause losses, and insurers and policyholders alike would have a better understanding of the scope of insurance policies. In a recent conversation with one of the nation's premier property insurance coverage experts, James Costner indicated that it is virtually impossible to maintain contract certainty in the current state of multi-cause loss jurisprudence.¹⁶⁸ Adopting the approach advocated in this article would undoubtedly improve contract certainty.

Second, most policyholders—personal lines and small commercial accounts—do not possess the power or capacity to unilaterally alter

¹⁶⁷ Bragg, *supra* note 120.

¹⁶⁸ Interview with James Costner, Former Senior Vice President, Property Practice, Willis North America (Feb. 9, 2011).

Insurance Services Office policies.¹⁶⁹ While courts often discuss the freedom to contract and reason that policyholders could have bargained to avoid the “anti-concurrent exclusions,” these courts fail to acknowledge that insurance contracts are pure contracts of adhesion, meaning that the contracts are presented on a take-it-or leave it basis. Additionally, the policyholder generally has no idea that anti-concurrent exclusions exist, much less any clue as to how the policies will be interpreted. Accordingly, the approach advocated in this article attempts to align doctrines of reasonableness with policyholder expectations. It should be unconscionable for a loss that is 99% covered to be excluded merely because 1% of the loss was excluded.

The unconscionability extends even further when the potential for a 99%-covered loss is excluded under an “**all risk**” policy. Policyholders understandably overestimate what is included in an “all risk” policy, but no reasonable policyholder would expect the disproportionate forfeiture that would result when a 99%-covered loss is excluded simply because a crafty adjuster is able to find some small amount of the loss that is excluded.¹⁷⁰ Also, if these types of exclusions are included, they should come with a disclaimer specifically alerting the policyholder of the nature of the potential exclusion.

Thus, while the approach advocated in this article does follow some elements of the dominant-cause approach, it is a deliberate choice, which more accurately reflects what should be the default position between insurer and policyholder.

This approach also limits the dominant-cause approach analysis to chain-of-event losses, and parts ways with the dominant-cause approach for independently caused losses. As advocated in this article, independent losses should be analyzed separately and distinctly from chain-of-event losses.¹⁷¹ Accordingly, the approach advocated in this article, attempts to provide a new method of analysis for independent losses.

¹⁶⁹ To be sure, large commercial entities can dictate terms of insurance and can negotiate, draft, and use manuscript policies.

¹⁷⁰ The policies themselves can also be drafted to ensure that some portion of the loss will be excluded. *See, e.g.,* Lexington Ins. Co. v. Unity/Waterford-Fair Oaks, Ltd., No. CIV.A. 399CV1623D, 2002 WL 356756, at *4 (N.D. Tex. Mar. 5, 2002) (denying coverage for improper “property maintenance” as a concurrent cause).

¹⁷¹ This idea is not novel and was discussed by the California Supreme Court in *Garvey*.

2. Confusion About How to Approach the Analysis

Because this approach recommends a series of questions relating to categorization, some may argue that the categorization itself could prove more problematic than the original problem. For instance, questions may arise as to how a court should determine whether causes harmonized to create a clear chain-of-events or whether the causes operated independently.

There certainly exists an element of discretion in this approach. There will invariably be close cases in determining whether a loss was caused by a chain-of-events or independent causes. Similarly, it may not always be easy to determine whether causes operated simultaneously or sequentially.

This approach addresses those concerns by making those close calls fact issues. The fact-finder will determine whether the causes operated sequentially or simultaneously. The approach is not designed to remove fact finding from the calculus. Rather, the approach attempts to clearly delineate fact questions from legal questions. Once the fact-finder determines the relevant facts, the law is more easily applied.

Certainly, there will be results where parties disagree with courts' conclusions respecting whether the losses were harmonious or independent. However, the facts will be uniformly applied and will generate some consistency in the muddled "concurrent causation" web. Further, the approach will allow courts to look to other jurisdictions and clearly understand how a court ruled and why the court ruled as it did.

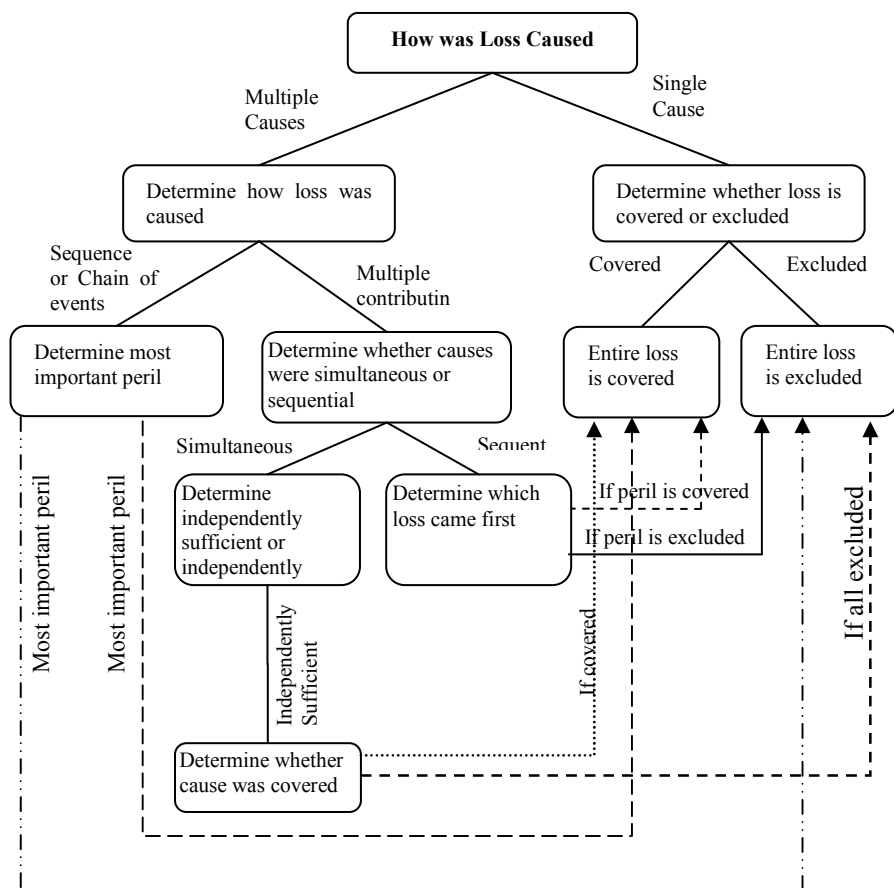
Creating clear legal guidelines will allow parties to understand *ex ante* the types of issues that will be addressed. Policyholders will have a clearer understanding of perils that are covered and excluded and will not have to play the concurrent-causation-roulette currently employed across jurisdictions. Similarly, insurers will understand how courts interpret their policies, which will create greater contract certainty and more accurate underwriting determinations.

VI. CONCLUSION

"Current causation" has evolved into an unworkable mess. The concurrent causation lexicon has become so muddled and amalgamated that it is impossible to forecast how a court, insurer, or policyholder will interpret "concurrent causation" questions. For these reasons, this article concludes that the "concurrent causation" lexicon should be revised and recommends that courts analyze multi-cause losses according to a

formulaic, categorical approach. By applying more precise and accurate language to multi-cause losses, courts and commentators will avoid unnecessary confusion and potential conflation of terms; thereby assuring contract certainty and ensuring that reasonable expectations are maintained.

APPENDIX



INSURANCE RATES REGULATION IN COMPARISON WITH OPEN COMPETITION

ANGELO BORSELLI*

TABLE OF CONTENTS

INTRODUCTION.....	112
I. HISTORICAL DEVELOPMENT OF RATE REGULATION.....	113
A. PRIVATE CONTROLS OVER INSURANCE RATES AND THE ENACTMENT OF ANTI-COMPACT LAWS.....	113
B. BEGINNING OF RATE REGULATION.....	115
C. <i>U.S. v. SOUTH-EASTERN UNDERWRITERS ASSOCIATION</i> AND THE MCCARRAN-FERGUSON ACT.....	117
D. THE NAIC-AIC MODEL RATE REGULATORY BILLS AND THE SUBSEQUENT STATE LEGISLATIVE ACTION.....	119
E. THE AFTERMATH OF THE ENACTMENT OF THE RATE REGULATORY BILLS.....	120
F. REVISION OF THE RATING LAWS.....	122
G. RECENT DEVELOPMENTS.....	124
II. RATIONALES FOR AND AGAINST RATE REGULATION.....	128
A. PURPOSES OF RATE REGULATION.....	128
B. ARGUMENTS FOR RATE REGULATION.....	131
C. ARGUMENTS AGAINST RATE REGULATION.....	132
III. THE EFFECTS OF OPEN COMPETITION.....	136
A. THE STRUCTURE OF THE U.S. PROPERTY AND CASUALTY INSURANCE MARKET.....	136
B. THE EUROPEAN EXPERIENCE WITH REGARD TO REGULATION OF INSURANCE TARIFFS.....	142
C. THE CASE FOR RATE DEREGULATION.....	151
CONCLUSION.....	160

* PhD Candidate in Business and Commercial Law, Università Bocconi, Italy. Insurance Law LL.M., University of Connecticut School of Law, 2009, *with honors*. Degree in law, Università della Calabria, Italy, 2008, *summa cum laude and distinction*.

This article examines rate regulation in the property and casualty insurance market in the United States. While rate regulation serves, in particular, the purpose of promoting insurer solvency and preventing oligopolistic pricing, it can also lead to market inefficiencies. The article argues for rate deregulation as a superior alternative to the current regulation model.

During the nineteenth century the property and casualty insurance market was highly competitive, featuring periods of low losses and large profits that attracted new market entrants. This competition caused insurance companies to set rates that were inadequate, thus leading to thousands of insolvencies. One method to solve the problem was the compact, an agreement among insurers to have a manager set rates. This solution often failed because members of the compact often cheated and there was no way to make every insurer in the market join. By the end of the nineteenth century several states had passed anti-compact law prohibiting the practice.

States in the early twentieth century started passing rate regulation laws in the fire insurance market. These laws subjected rate setting to state control. These laws were prevalent in the fire insurance market by the 1940's, but were not widespread in the casualty insurance market. With the passage of the McCarran-Ferguson Act in 1945, insurance regulation was made the primary purview of the several states, with federal intervention allowable only where states failed to legislate. Soon after, the AIC and NAIC crafted model laws that became the basis for much state legislation. While state laws started out with an emphasis on cooperative rate setting by rate bureaus, gradually the laws were changed to facilitate independent rate filing by insurers directly to state insurance regulators, thus increasing rate competition. Today, the trend is toward less restrictive systems of rate regulation in most property lines: gradually away from "prior approval" towards "file and use," "use and file," "flex rating," "modified prior approval," or no file systems.

The rationales for rate regulation include consumer protection, the prevention of insurer insolvency and unfair pricing, and the promotion of actuarial accuracy. The rationales against rate regulation are mainly that cartel pricing and destructive competition are not a threat today, that the market is the appropriate price setting mechanism in insurance markets, deregulation promotes competition, and that rate deregulation would take the politics out of rate setting.

Examining the structure of the American property and casualty insurance market is necessary to determine how successful a policy of deregulation will be. The U.S. property and casualty insurance market

presents the structural features of a competitive market because it is characterized by a large number of firms selling products with identical features. Evaluation of the Herfindahl-Hirschman Index for industry concentration in conjunction with the Department of Justice's Merger Guidelines indicates that the property and casualty insurance market is not concentrated. Furthermore, the trend toward an increase in the level of market concentration is oftentimes the result of market competition since it leads to low-cost, efficient firms replacing high-cost firms. The property and casualty insurance market is also widely regarded as having low barriers to entry for new firms. Thus the market does not have monopolistic or oligopolistic characteristics that would justify rate regulation. Nevertheless, purely competitive rate setting systems are seldom used throughout the United States.

The European Union can provide a helpful case study in considering rate setting deregulation because EU member states do not have the right to regulate insurance prices, after the European Parliament passed the third non-life insurance Directive in 1992. Previously, EU member states exercised considerable rate setting power. The experience has been a positive one. Competition increased, especially in heavily regulated markets, and premium rates decreased. Market concentration, however, did not decrease, though this could be attributed to an increase in mergers and acquisitions. The number of insolvencies decreased as prices were better aligned with costs.

Rate regulation in the United States may be adversely impacting insurer profitability, as rate changes are impeded as market conditions change. There is empirical evidence that property and casualty insurance companies have experienced a lower rate of return than other industries. These artificially low returns may have led to many insurers' exits from the market. In particular, with regard to some lines, over the 2000-2009 period, more insurers exited the market than entered it. Deregulation would eliminate compliance costs and allow rate changes. Even if rate deregulation lead to higher rates, in the long run this would be offset by greater market availability and consumer choice. Rate regulation has the tendency to force stricter underwriting that limits market availability. Additionally, there is no evidence that rate regulation has eliminated the possibility of insurer insolvency. It is more likely that allowing insurance companies to set rates commensurate with their costs will enhance their financial strength. Therefore policy makers should seriously consider greater insurance rate deregulation.

INTRODUCTION

The traditional justification for economic regulation is to protect the public interest by correcting market failures and improving economic efficiency and equity.¹ In particular, with regard to insurance, regulation aims to protect policyholders by ensuring the solvency of the insurance companies. In light of the peculiar nature of an insurance contract, in which the policyholder pays an upfront premium in exchange for the insurer's promise to pay in case a loss occurs, the need is clear to assure the financial solidity of the insurance companies and their ability to pay possible future claims.

In this context, by the first half of the twentieth century individual states within the United States enacted rate regulatory laws to ensure that rates were "adequate, not excessive and not unfairly discriminatory." One objective of rate regulation is to prevent insolvencies by avoiding a sort of ruinous competition in which insurers, in order to strengthen their market position, charge rates not sufficient to cover their costs. Another is to avoid oligopolistic pricing.

This study will focus on the regulation of rates in the U.S. property and casualty insurance market, highlighting the inefficiencies caused by the system. The aim of the paper is to examine the advisability of replacing rate regulation with rate deregulation. In this regard, although some states have rating methods less restrictive than prior approval, like file and use, use and file, flex rating and modified prior approval, it must be emphasized that none of these methods fully rely on competition since the insurance commissioner basically still retains the right to direct the insurers' setting of rates. The analysis supports the conclusion that rate deregulation should be introduced.

Part I will provide the historical background of rate regulation, discussing the developments from the nineteenth century, when rate regulation was introduced in order to prevent insurers' insolvencies, to the more recent trend toward less restrictive rating laws.

Part II will set out the purposes of rate regulation to ensure, as stated above, that insurance rates are "adequate, not excessive and not

¹ STEPHEN BREYER, REGULATION AND ITS REFORM 15-16 (1982); SCOTT E. HARRINGTON, INSURANCE DEREGULATION AND THE PUBLIC INTEREST 15 (2000); 2 ALFRED E. KAHN, THE ECONOMICS OF REGULATION, INSTITUTIONAL ISSUES 11 (1988).

unfairly discriminatory.” Further, the arguments adduced for and against rate regulation will be presented.

Part III will make the case for rate deregulation. In particular, the analysis will consider the U.S. property and casualty insurance market structure, the performance of the industry, market growth, market entries and exits and the effects of rate regulation on insurance availability. The analysis also considers the experience of the European Union, where state supervisory authorities have been prevented from exerting control over insurance premiums prices.

I. HISTORICAL DEVELOPMENT OF RATE REGULATION

A. PRIVATE CONTROLS OVER INSURANCE RATES AND THE ENACTMENT OF ANTI-COMPACT LAWS

During the nineteenth century, the property and casualty insurance market was distinguished by a high level of competition. Indeed, since historically the fire insurance business was highly cyclical, in periods when losses were low and profits high new insurance companies entered the market aiming to make large profits.² Neither barriers to entry nor significant economies of scale hindered the entrance into the market.³ The strong competition in the market in the 1800’s led insurers to set inadequate rates and thus, by 1877, around 3000 companies had become insolvent.⁴

In response, in 1866 insurers instituted a national organization, the National Board of Fire Underwriters, in order to “establish and maintain, as far as practicable, a system of uniform rates of premium.”⁵ However, the fact that, in profitable periods, insurance companies violated the agreements concerning the rates established by the Board, made the Board

² 1 JON S. HANSON ET AL., NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS, MONITORING COMPETITION: A MEANS OF REGULATING THE PROPERTY AND LIABILITY INSURANCE BUSINESS 9 (1974); Spencer L. Kimball & Ronald N. Boyce, *The Adequacy of State Insurance Rate Regulation: The McCarran-Ferguson Act in Historical Perspective*, 56 MICH. L. REV. 545, 547 (1958); Michael D. Rose, *State Regulation of Property and Casualty Insurance Rates*, 28 OHIO ST. L.J. 669, 677 (1967).

³ Rose, *supra* note 2, at 677.

⁴ HANSON ET AL., *supra* note 2, at 9; Kimball & Boyce, *supra* note 2, at 547-48; Rose, *supra* note 2, at 677.

⁵ Rose, *supra* note 2, at 677 (quoting Kimball & Boyce, *supra* note 2, at 548).

ineffective in controlling rates.⁶ Nevertheless, after the fire losses that occurred in Chicago and Boston respectively in 1871 and 1872, it appeared inevitable that insurance companies had to cooperate in order to set adequate rates.⁷ In 1877, the National Board of Fire Underwriters abandoned its function of rate control – addressing itself only to fire prevention and the maintenance of statistics – and in the 1880s, regional associations of companies assumed the task of rate stabilization.⁸ Among the techniques implemented by the regional associations in order to control rates was the compact, an agreement according to which the compact manager set the rates and usually enforced them in compliance with the compact's conditions.⁹

However, the problem any cartel faces is that each cartel member has incentives to raise its profits by cheating the cartel: bringing down the cartel's price and increasing its output.¹⁰ In the same way, these regional associations did not effectively stabilize insurance rates.¹¹ Indeed, the insurer members of the association did not always honor the agreements made in good faith since they used to cut rates.¹² Further, the agreements were often undermined by insurers that were not members of the association.¹³

Toward the end of the 1800's, many states responded to the insurers' efforts to fix rates by passing anti-compact laws.¹⁴ The first anti-compact laws were passed by Ohio and New Hampshire in 1885¹⁵ and by 1912 twenty-three states had passed such type of legislation.¹⁶

⁶ HANSON ET AL., *supra* note 2, at 10-11; Kimball & Boyce, *supra* note 2, at 548; Rose, *supra* note 2, at 677.

⁷ Kimball & Boyce, *supra* note 2, at 548.

⁸ HANSON ET AL., *supra* note 2, at 11; Kimball & Boyce, *supra* note 2, at 549; Kent H. Parker, *Ratemaking in Fire Insurance*, in PROPERTY AND LIABILITY INSURANCE HANDBOOK 169, 170 (John D. Long & Davis W. Gregg eds., 1965); Rose, *supra* note 2, at 677-78.

⁹ HANSON ET AL., *supra* note 2, at 11; Kimball & Boyce, *supra* note 2, at 549. The first compact was the St. Louis Compact, concluded in 1879.

¹⁰ LEE S. FRIEDMAN, THE MICROECONOMICS OF PUBLIC POLICY ANALYSIS 668 (2002); DAVID A. BESANKO & RONALD R. BRAEUTIGAM, MICROECONOMICS 495 (2002); *see generally* D. K. Osborne, *Cartel Problems*, 66 AM. ECON. REV. 835-44 (1976).

¹¹ HANSON ET AL., *supra* note 2, at 12-13.

¹² *Id.* at 12.

¹³ *Id.*

¹⁴ Kimball & Boyce, *supra* note 2, at 549; Rose, *supra* note 2, at 678.

¹⁵ HANSON ET AL., *supra* note 2, at 14.

¹⁶ Kimball & Boyce, *supra* note 2, at 550.

Nevertheless, the anti-compact laws were often evaded. In cases where the law expressly prohibited agreements between insurance companies, like in Ohio and Wisconsin, insurers concluded that agreements between agents were not prohibited. So, the insurance companies relied on their agents to fix rates.¹⁷ Elsewhere, where the law prohibited all agreements relating to the establishment of rates, insurance companies formed “independent” bureaus to make advisory rates.¹⁸

B. BEGINNING OF RATE REGULATION

During the first half of the twentieth century, states became aware of the risks that ruinous competition posed to policyholders and began to enact legislation to regulate fire insurance rates.¹⁹ The first rate regulatory law, passed in Kansas in 1909, required fire insurers to file their rates and their rating plans with the superintendent of insurance and prohibited rate discrimination among insureds.²⁰ It also required insurance companies to give the insurance commissioner ten days’ notice in order to change rates and authorized the commissioner to adjust rates that were excessive or inadequate.²¹ In 1914, the German Alliance Insurance Company challenged the Kansas law as unconstitutional.²² The company argued that insurance is a private contract and that the state has no power to interfere by regulating insurance rates; otherwise, such regulation would be a deprivation of the insurer’s property without due process of law, in violation of the Fourteenth Amendment to the Constitution.²³ The Supreme Court rejected the complaint, stating that insurance was affected with a public interest and, for this reason, the insurance premium could be fixed by law.²⁴

After the San Francisco fire of 1909, the New York legislature conducted an investigation on fire insurance rating practices.²⁵ To this end a Joint Legislative Committee, known as the Merritt Committee, was

¹⁷ Rose, *supra* note 2, at 678.

¹⁸ Rose, *supra* note 2, at 678; Richard A. Wiley, *Pups, Plants and Package Policies – or the Insurance Antitrust Exemption Re-Examined*, 6 VILL. L. REV. 281, 312-14 (1961).

¹⁹ Kimball & Boyce, *supra* note 2, at 551; Rose, *supra* note 2, at 679.

²⁰ Kimball & Boyce, *supra* note 2, at 551; Rose, *supra* note 2, at 679.

²¹ Kimball & Boyce, *supra* note 2, at 551; Rose, *supra* note 2, at 679.

²² German Alliance Ins. Co. v. Lewis, 233 U.S. 389, 389 (1914).

²³ *Id.* at 397.

²⁴ *Id.* at 414-18. See also Nat’l Union Fire Ins. Co. v. Wanberg, 260 U.S. 71, 75 (1922) (holding that insurance is a business affected with a public interest).

²⁵ HANSON ET AL., *supra* note 2, at 17; Rose, *supra* note 2, at 679-80.

established.²⁶ The Committee did not view the results of the anti-compact laws positively since they led to destructive competition and rate discrimination.²⁷ It instead recommended the passage of a statute providing for the filing by rate bureaus of fire insurance rates with the Insurance Department and subjecting those bureaus to the Insurance Department's control.²⁸ In accordance with these recommendations, the New York legislature enacted a law allowing fire insurers to fix rates in concert.²⁹ Rate bureaus were authorized and had to set the rates; the rates had to be filed with the insurance superintendent, who had to approve them before they could be used.³⁰

Afterwards other states, acknowledging the inefficiency of the anti-compact laws, passed similar rate regulatory laws.³¹ States no longer relied on competition as a means for rate-setting; instead, they authorized rating bureaus to set fire insurance rates. Rate bureaus evolved from both the public and the industry interest in rate setting; they were privately operated except in Texas.³² Some states required companies "to become a member of or subscriber to a rating organization."³³ Rates were, however, subject to the public control by state insurance departments, which usually had to approve them. By 1944, there were only three states with no public control of rate-setting.³⁴

Up to 1945 states did not regulate rates for the casualty insurance industry to the same extent as in fire insurance.³⁵ Except for workmen's compensation insurance, most regulation aimed at avoiding rate discrimination.³⁶ Further, only a small number of states required filing and approval of automobile insurance rates.³⁷ In general, rate competition was prevalent in the casualty insurance market.³⁸

²⁶ HANSON ET AL., *supra* note 2, at 17; Rose, *supra* note 2, at 679-80.

²⁷ HANSON ET AL., *supra* note 2, at 17; Rose, *supra* note 2, at 679-80.

²⁸ HANSON ET AL., *supra* note 2, at 19; Rose, *supra* note 2, at 680.

²⁹ HANSON ET AL., *supra* note 2, at 19; Rose, *supra* note 2, at 680.

³⁰ Rose, *supra* note 2, at 680.

³¹ HANSON ET AL., *supra* note 2, at 19.

³² HANSON ET AL., *supra* note 2, at 20; Kimball & Boyce, *supra* note 2, at 552; Wiley, *supra* note 18, at 314.

³³ HANSON ET AL., *supra* note 2, at 20.

³⁴ Kimball & Boyce, *supra* note 2, at 551.

³⁵ HANSON ET AL., *supra* note 2, at 20; Rose, *supra* note 2, at 682.

³⁶ HANSON ET AL., *supra* note 2, at 20-21; Rose, *supra* note 2, at 682.

³⁷ HANSON ET AL., *supra* note 2, at 21; Rose, *supra* note 2, at 682.

³⁸ Rose, *supra* note 2, at 682.

C. *U.S. v. SOUTH-EASTERN UNDERWRITERS ASSOCIATION* AND THE
MCCARRAN-FERGUSON ACT

Until 1944, the U.S. Supreme Court's 1868 holding in *Paul v. Virginia* exempted rate-setting agreements from federal antitrust law.³⁹ The case involved an 1866 Virginia statute that prohibited insurers who were not incorporated in Virginia and their local agents from doing business in the state without first obtaining a license. The statute was challenged on the ground, *inter alia*, that it conflicted with the Commerce Clause of the U.S. Constitution,⁴⁰ which gives Congress the power "to regulate commerce with foreign nations, and among the several States."⁴¹ The Court, upholding the Virginia statute, held that insurance was not commerce, interstate or otherwise.⁴² From then on, insurance contracts were not subject to federal antitrust law.⁴³

In 1944, however, in *U.S. v. South-Eastern Underwriters Association*,⁴⁴ the Supreme Court reversed *Paul v. Virginia*. In that case, 198 member companies of the South-Eastern Underwriters Association and twenty-seven individuals were indicted in the U.S. District Court for the Northern District of Georgia.⁴⁵ The indictment alleged two conspiracies in violation of the Sherman Act.⁴⁶ The first was a conspiracy to restrain interstate commerce by fixing insurance premiums.⁴⁷ The second was a conspiracy to monopolize trade and commerce in the fire insurance sector and in allied lines in the states of Alabama, Florida, Georgia, North Carolina, South Carolina and Virginia.⁴⁸ The District Court dismissed the indictment and, relying on *Paul v. Virginia*, held that insurance was not commerce and therefore price-fixing in the business of insurance did not violate the Sherman Act.⁴⁹ On appeal, the Supreme Court ruled that the Sherman Act did apply to the fire insurance business since any business conducted across state lines was "commerce among the several States."⁵⁰ In

³⁹ *Paul v. Virginia*, 75 U.S. 168, 183 (1868).

⁴⁰ *Id.* at 177.

⁴¹ U.S. CONST. art. I, § 8, cl. 3.

⁴² *Paul*, 75 U.S. at 183.

⁴³ *Id.*

⁴⁴ *U.S. v. Se. Underwriters Ass'n*, 51 F. Supp. 712, 713 (N.D. Ga. 1943).

⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ *Id.*

⁴⁸ *Id.*

⁴⁹ *Id.* at 713-15.

⁵⁰ *U.S. v. Se. Underwriters Ass'n*, 322 U.S. 533, 538-41 (1944).

that regard, the Court specified that all commercial activities conducted across state lines fell within Congress' regulatory power under the Commerce Clause and no exception could be made for the business of insurance.⁵¹

Following this decision, insurance companies and states lobbied Congress to avoid federal regulation of the insurance sector.⁵² In particular, states were afraid to lose their regulatory power and the power to tax insurance companies.⁵³ In 1945 Congress passed the McCarran-Ferguson Act to preserve the states' control over insurance regulation.⁵⁴ Congress was concerned about the uncertainty that might ensue from a change in regulatory authority and also believed that states could regulate insurance better than the federal government, because of their relationship with the insurance companies and their experience with regulating insurance.⁵⁵

In the preamble McCarran-Ferguson states that "the continued regulation and taxation by the several States of the business of insurance is in the public interest."⁵⁶ To that end, Section 1012(b) of McCarran-Ferguson provides that no "Act of Congress shall be construed to invalidate, impair or supersede" any state law enacted in order to regulate or tax the business of insurance.⁵⁷ By virtue of this provision, state law supplanted federal antitrust regulation of the insurance industry. Conversely, the Act provides that the Sherman, Clayton and Federal Trade Commission Acts apply to "the business of insurance to the extent that such business is not regulated by State law,"⁵⁸ except for agreements or acts of boycott, coercion, or intimidation.⁵⁹ This aspect of the act was considered a compromise between those in Congress who favored an antitrust exemption and those who favored federal supervision of the insurance industry.⁶⁰ In this way the act permits states the opportunity to continue to regulate insurance, while retaining the right for federal intervention in case the states failed to intervene.⁶¹ As a consequence of

⁵¹ *Id.* at 553.

⁵² Kimball & Boyce, *supra* note 2, at 553-54.

⁵³ *Id.* at 554.

⁵⁴ 15 U.S.C.A. §§ 1011-1015 (West 2009).

⁵⁵ Rose, *supra* note 2, at 694.

⁵⁶ 15 U.S.C.A. § 1011 (West 2009).

⁵⁷ *Id.* § 1012(b).

⁵⁸ *Id.*

⁵⁹ *Id.* § 1013(b).

⁶⁰ Kimball & Boyce, *supra* note 2, at 555.

⁶¹ *Id.*

McCarran-Ferguson most states enacted laws regulating insurance and, in particular, rate-setting.⁶²

D. THE NAIC-AIC MODEL RATE REGULATORY BILLS AND THE
SUBSEQUENT STATE LEGISLATIVE ACTION

Section 1012(b) of McCarran-Ferguson encouraged the states to regulate insurance in order to avoid federal intervention.⁶³ Soon, uniform legislation was introduced in the states under the auspices of the NAIC and the All-Industry Committee (AIC).⁶⁴ The AIC, representing nineteen insurer associations, was formed to cooperate with the NAIC to develop model legislation designed to reinforce state control of insurance in accordance with section 1012(b).⁶⁵ In 1946, two model laws, one for fire, marine and inland marine insurance and the other for casualty and surety insurance, were submitted to individual states for passage.⁶⁶

The two model laws, which were similar in content, proposed proscriptions on excessive, inadequate or unfairly discriminatory rates and advocated supervision of rate-setting among insurers.⁶⁷ The bills provided that in setting rates, consideration should be given, *inter alia*, to past and prospective loss experience in the state and elsewhere.⁶⁸ Further, insurers had to file any rating plan and any modification to that plan with state insurance commissioners.⁶⁹ Under the model laws, such information would become public after the filing became effective.⁷⁰

The model laws also addressed rate-setting by insurers. Insurers were allowed to benefit from the services of rating organizations or of advisory organizations.⁷¹ Rating organizations made rates for their members and subscribers, while advisory organizations assisted insurers which filed their own rates or rating organizations in rate making by

⁶² Rose, *supra* note 2, at 696.

⁶³ *Id.* (quoting Patrick A. McCarran, *Federal Control of Insurance: Moratorium under Public Law 15 Expired July 1*, 34 A.B.A. J. 539, 540 (1948)).

⁶⁴ Kimball & Boyce, *supra* note 2, at 555; Rose, *supra* note 2, at 696-97.

⁶⁵ Rose, *supra* note 2, at 697.

⁶⁶ Kimball & Boyce, *supra* note 2, at 555; Rose, *supra* note 2, at 698-99.

⁶⁷ HANSON ET AL., *supra* note 2, at 29-33 (reproducing the draft of the casualty and surety insurance bill); Rose, *supra* note 2, at 699-701 (reproducing the draft of the fire, marine and inland marine insurance bill).

⁶⁸ HANSON ET AL., *supra* note 2, at 30; Rose, *supra* note 2, at 699.

⁶⁹ HANSON ET AL., *supra* note 2, at 30; Rose, *supra* note 2, at 699.

⁷⁰ HANSON ET AL., *supra* note 2, at 30; Rose, *supra* note 2, at 699.

⁷¹ HANSON ET AL., *supra* note 2, at 31-32; Rose, *supra* note 2, at 699-700.

collecting and furnishing loss and expense data or by providing recommendations concerning rates.⁷² Both rating organizations and advisory organizations were subject to state requirements.⁷³ Insurers could file independent rates or those made by a licensed rating organization. Insurers were also allowed to seek permission from commissioners to file deviations from rates set by a rating organization.⁷⁴ Lastly, under the so-called “deemer clause”, rate filings were considered approved unless disapproved within fifteen days, or thirty days if the commissioner decided to extend the period for approval.⁷⁵

The model bills, which favored the interests of the rate bureaus,⁷⁶ were introduced with amendments in some states.⁷⁷ In particular, the amendments concerned the deemer clause since states had different interpretations about how rates should be filed with commissioners in order to meet the state regulation requirement of section 1012(b) of the McCarran-Ferguson Act.⁷⁸ Some states, like California, Missouri and Idaho, did not require rate filings for fire or casualty lines (although the insurance commissioner had discretionary authority to request such filings), while other states, like Delaware, Maine, Massachusetts, Wyoming, Ohio and District of Columbia, provided that the rates once filed became effective, subject to subsequent disapproval.⁷⁹ The model laws assumed that the requirements for reverse preemption found in section 1012(b) of McCarran-Ferguson were met by the mere existence of state legislation.⁸⁰

E. THE AFTERMATH OF THE ENACTMENT OF THE RATE REGULATORY BILLS

With the enactment of state rate regulatory laws based on the NAIC-AIC model bills, several controversies about competitive versus cooperative rate making arose.⁸¹ Between 1947 and 1957 the rate bureaus

⁷² HANSON ET AL., *supra* note 2, at 31-32; Rose, *supra* note 2, at 700.

⁷³ HANSON ET AL., *supra* note 2, at 31-32; Rose, *supra* note 2, at 700.

⁷⁴ HANSON ET AL., *supra* note 2, at 30-31; Rose, *supra* note 2, at 700.

⁷⁵ HANSON ET AL., *supra* note 2, at 30; Rose, *supra* note 2, at 701.

⁷⁶ Rose, *supra* note 2, at 698, 703.

⁷⁷ Kimball & Boyce, *supra* note 2, at 555.

⁷⁸ Rose, *supra* note 2, at 704.

⁷⁹ *Id.* at 704 & nn.176-77.

⁸⁰ See *id.* at 705-06 (describing the legislative history of the McCarran-Ferguson Act).

⁸¹ *Id.* at 716-17.

put up strong resistance to rate deviations.⁸² Insurers seeking to file deviations experienced obstacles because the NAIC-AIC model laws required commissioners to notify the rate bureaus before approving a rate deviation.⁸³ The rate bureaus could then testify in opposition to the deviation.⁸⁴ Further, since the deviation filing was valid for only one year, insurers had to, at considerable cost, justify their request annually.⁸⁵ This led many insurers to resign from the bureaus in order to make independent filings.⁸⁶ For example, in 1954 the Insurance Company of North America (INA) resigned from the New York Fire Insurance Rating Organization (NYFIRO) and made independent rate filings for most dwelling classes while remaining a subscriber for other dwelling classes and commercial lines.⁸⁷ The NYFIRO challenged the New York department's approval, arguing that INA could not independently file for some risks and subscribe to the bureau for others and that INA violated NYFIRO's property rights by using bureau data in its filings.⁸⁸ The New York Insurance Department rejected the NYFIRO's petition and authorized independent filing and the right of partial subscribership.⁸⁹

In 1955 the Pacific Fire Rate Bureau adopted a rule that companies that made independent filings could no longer benefit from bureau scheduled-rating services.⁹⁰ The rule was challenged in several states since insurers valued the scheduled-rating services and did not want to lose their right to subscribe to them.⁹¹ The Arizona Supreme Court held the rule invalid in 1958.⁹²

The partial subscribership system that resulted permitted the setting of more independent rates since insurance companies with sufficient loss experience in certain lines of insurance could make independent filings.⁹³ In doing so, insurers had to confront the attempts by the bureaus to

⁸² *See id.* at 718.

⁸³ *Id.*

⁸⁴ Rose, *supra* note 2, at 718.

⁸⁵ *Id.*

⁸⁶ *Id.* at 720.

⁸⁷ *Id.*

⁸⁸ *Id.* at 720-21.

⁸⁹ *Id.* at 721.

⁹⁰ Rose, *supra* note 2, at 721.

⁹¹ *Id.*

⁹² *Pac. Fire Rating Bureau v. Ins. Co. of N. Am.*, 321 P.2d 1030, 1034-35 (Ariz. 1958).

⁹³ Joel B. Dirlam & Irwin M. Stelzer, *The Insurance Industry: A Case Study in the Workability of Regulated Competition*, 107 U. PA. L. REV. 199, 208 (1959).

intervene in the hearings and oppose the deviations as aggrieved parties, subjecting insurers to delays and expenses.⁹⁴ To remedy this situation, insurance commissioners and courts⁹⁵ ruled that rate bureaus could not appear as aggrieved parties because they acted not to benefit the public but rather to protect their own interests.⁹⁶ In the process, a more flexible rate setting system emerged that permitted price competition to a certain extent through deviations from the bureau rates, aided by provisions in the NAIC-AIC model laws that did not make membership in rate bureaus mandatory, allowing insurers to make independent filings.

F. REVISION OF THE RATING LAWS

The trend toward less regulated rates was also reflected in a study conducted by the Senate Antitrust and Monopoly Subcommittee in 1958. The Senate appointed the Subcommittee to conduct a study on insurance and antitrust laws. The Subcommittee recommended denying rate bureaus the status of aggrieved parties, eliminating the requirement of the annual filings of deviations and, lastly, adopting file-and-use rating systems.⁹⁷

Similar recommendations were also provided by the subcommittee appointed by the NAIC Rates and Rating Organization Committee in 1960 in order to review the insurance state regulation system.⁹⁸ The subcommittee recommended that no rate bureau should have the status of aggrieved party because the bureaus had no interest in decisions on rate filings. Rate bureaus, according to the subcommittee, should have been denied status to apply for a hearing on insurers' independent filings and the one-year limitation on deviation should have been eliminated. The subcommittee also recommended continuing the deemer clause and the right of partial subscribership to bureaus and to consolidate the fire and casualty rating bills in order to permit the development of multi-line package policies.⁹⁹

In 1962 the NAIC approved amendments to the model rating laws, adopting the recommendations of its subcommittee.¹⁰⁰ In 1964 the NAIC

⁹⁴ See Rose, *supra* note 2, at 723-24.

⁹⁵ See, e.g., Va. Ass'n of Ins. Agents v. Commonwealth, 110 S.E.2d 223, 228 (Va. 1959) (holding that rate bureaus could not be granted the status of aggrieved party since orders concerning rates were not addressed to them).

⁹⁶ Rose, *supra* note 2, at 725.

⁹⁷ *Id.* at 725-26.

⁹⁸ HANSON ET AL., *supra* note 2, at 46-47.

⁹⁹ *Id.* at 46-47.

¹⁰⁰ HANSON ET AL., *supra* note 2, at 48; Rose, *supra* note 2, at 726.

Rates and Rating Organizations Committee appointed another subcommittee to inquire into the rate regulation system.¹⁰¹ The subcommittee found that competition had increased in the insurance market since the 1940s.¹⁰² In particular, it noted that price competition had become more widespread due to independent rate filings and decreased influence by rate bureaus in setting rates.¹⁰³ The Subcommittee recommended placing more reliance on fair competition to set insurance prices¹⁰⁴ and suggested the suspension where possible of the prior approval system and its replacement with no prior approval rate regulation.¹⁰⁵ In states where local market characteristics did not permit such a change, the subcommittee recommended continuing the deemer provision to assure prompt responses to rate filings.¹⁰⁶

In a file and use system a rate filing becomes effective once the rate is filed with the insurance commissioner. Soon, the insurance industry embraced no prior rate approval and file and use provisions to permit insurers to respond immediately to market changes.¹⁰⁷ By 1985, 24 states adopted such changes in their rating laws.¹⁰⁸

California, for example, adopted a competitive rate setting system that incorporated a no filing provision and abolished any requirement to belong to a rate bureau.¹⁰⁹ Under California law, rates could be used immediately without having to be filed or approved by the commissioner. Illinois, which had originally enacted a prior approval law, enacted an open competition law in 1970.¹¹⁰ The open competition rating law was distinguished by the lack of advance approval or disapproval by the

¹⁰¹ HANSON ET AL., *supra* note 2, at 48.

¹⁰² *Id.* at 51.

¹⁰³ *Id.*

¹⁰⁴ *Id.*

¹⁰⁵ *Id.*

¹⁰⁶ *Id.* at 53.

¹⁰⁷ Rose, *supra* note 2, at 726-28.

¹⁰⁸ Paul L. Joskow & Linda McLaughlin, *McCarran-Ferguson Act Reform: More Competition or More Regulation?*, 4 J. RISK & UNCERTAINTY 373, 380 (1991).

¹⁰⁹ CAL. INS. CODE §§ 1850-1850.3 (1947) (repealed 1988); HANSON ET AL., *supra* note 2, at 56, 395-419.

¹¹⁰ HANSON ET AL., *supra* note 2, at 57, 420; Stephen P. D'Arcy, *Insurance Price Deregulation: The Illinois Experience*, in DEREGULATING PROPERTY-LIABILITY INSURANCE: RESTORING COMPETITION AND INCREASING MARKET EFFICIENCY 248, 256 (David J. Cummins ed., 2002).

regulator and the prohibition of any agreement to adhere to bureau rates.¹¹¹ In August 1971¹¹² the law was allowed to sunset and Illinois became the only state without a rating law for property and casualty insurance.¹¹³ However, there has been no attempt at federal antitrust enforcement in Illinois since the two largest personal lines insurers domiciled there, State Farm and Allstate, make their rates independently.¹¹⁴ In June 1972 the state enacted a law limited to regulating advisory organizations which were defined to mean every person, other than an insurer, who compiles insurance statistics, prepares insurance policies and underwriting rules, makes surveys and insurance research and furnishes that material to insurance companies.¹¹⁵ Insurers were prohibited from agreeing with each other or the advisory organization to adhere to the use of any statistics, policy forms or underwriting rules.¹¹⁶

G. RECENT DEVELOPMENTS

Between 1985 and 1988 several states adopted flex-rating systems, seeking to establish caps on price increases.¹¹⁷ Indeed, especially in the area of auto insurance, regulators started focusing on affordability of coverage and suppression of rates despite increasing claim costs.¹¹⁸ In this connection, an important regulatory development occurred in California on November 8, 1988 with the passage of Proposition 103.¹¹⁹ Proposition 103 required a rollback of rates for automobile insurance to 20 percent below the rates in effect on November 8, 1987¹²⁰ unless the downturn in rates would have led to the insolvency of the insurer.¹²¹ In addition, a prior

¹¹¹ HANSON ET AL., *supra* note 2, at 395.

¹¹² HANSON ET AL., *supra* note 2, at 57, 420; D'Arcy, *supra* note 110, at 257.

¹¹³ D'Arcy, *supra* note 110, at 257.

¹¹⁴ *Id.*

¹¹⁵ HANSON ET AL., *supra* note 2, at 57, 420-21.

¹¹⁶ *Id.* at 57, 421.

¹¹⁷ Joskow & McLaughlin, *supra* note 108, at 380.

¹¹⁸ Scott E. Harrington, *Rate Suppression*, 59 J. RISK & INS. 185, 187 (1992) (defining "rate suppression" as government suppression of insurance rates below the level that would exist in the absence of price regulation).

¹¹⁹ CAL. INS. CODE §§ 1861.01-1861.02 (2008).

¹²⁰ *Id.* § 1861.01(a).

¹²¹ *Id.* § 1861.01(b). Subdivision (b) of § 1861.01 has been held unconstitutional on May 4, 1989 in the case of *Califarm Ins. Co. v. Deukmejian*, 771 P.2d 1247, 1255 (Cal. 1989) (reaffirming the constitutional standard of a fair

approval system was introduced for most casualty insurance rates. Beginning on November 8, 1989, property and casualty insurance rates in California had to be approved by the commissioner prior to use.¹²² Finally, Proposition 103 provided that personal automobile insurance rates must be determined taking into account, in decreasing order of importance, the insured's driving safety record, the number of miles driven annually by the insured, the number of years of driving experience of the insured and any other rating factors that the commissioner specified had a substantial relationship to the risk of loss.¹²³ A mandatory 20 percent discount for good drivers was also established.¹²⁴ Thus, Proposition 103 replaced the open competition system in force until then in California.

In 1994, in *20th Century Ins. Co. v. Garamendi*, the California Supreme Court affirmed the Insurance Commissioner's authority to adopt a ratemaking formula implementing the rate rollback provision of Proposition 103.¹²⁵ The court held that the rate making formula was not confiscatory since it did not preclude the setting of a just and reasonable rate.¹²⁶ According to the court, the rates set under the formula did not inflict "deep financial hardship" on insurers and therefore did not prevent them from operating successfully.¹²⁷ Going forward, California's system of rate regulation mainly aimed to avoid excessive rates by determining maximum rate levels.¹²⁸

In 1980 the NAIC adopted model laws for less restrictive rating systems introducing the "file and use" and "use and file" types of rate regulation.¹²⁹ In the 1990s, catastrophic losses increased the level of state intervention in the pricing and underwriting of homeowners' insurance.¹³⁰ A tendency towards less restrictive rate regulation emerged with regard to other property lines, in particular automobile insurance.¹³¹ Favorable loss

and reasonable return according to the due process clause of the state and federal Constitution).

¹²² CAL. INS. CODE § 1861.01(c) (2008).

¹²³ *Id.* § 1861.02(a).

¹²⁴ *Id.* § 1861.02(b)(2).

¹²⁵ *20th Century Ins. Co. v. Garamendi*, 878 P.2d 566 (Cal. 1994).

¹²⁶ *Id.* at 617-18.

¹²⁷ *Id.*

¹²⁸ Harrington, *supra* note 118; Joskow & McLaughlin, *supra* note 108.

¹²⁹ NAIC, PERSONAL LINES REGULATORY FRAMEWORK 2 (draft white paper) (Apr. 16, 2009), www.naic.org/documents/committees_papers_framework_02.doc.

¹³⁰ HARRINGTON, *supra* note 1, at 9-10.

¹³¹ *Id.* at 10.

trends led insurers to reduce their rates and consequently the need to use regulation to suppress rates declined.¹³²

Toward the end of the 1990s, many states passed laws deregulating the price and policy forms of commercial insurance.¹³³ Commercial deregulation laws were enacted in 1998 in Arizona, Georgia, Illinois and New Hampshire and in 1999 in Arkansas, Colorado, Kansas, Montana, Missouri, Oklahoma, Indiana, Maine, Louisiana, Virginia and Rhode Island.¹³⁴ New York, Connecticut and Massachusetts have also adopted similar legislation.¹³⁵ These laws exempt insurance companies that sell their products to large specialized commercial insurance buyers from rate and policy form requirements.¹³⁶ The hope was that if insurers did not have to comply with state control, they would be able to diversify both their rates and types of policies, thereby expanding the range of products they could offer.¹³⁷

Currently, there are six different types of rate regulation systems.¹³⁸ Rate regulation varies from the most restrictive type, the prior approval method, to the no-filing method, the least restrictive. The six systems are the prior approval, file and use, use and file, flex rating, modified prior approval, and no file methods.¹³⁹ The prior approval system requires insurers to file the rates and wait for the approval by the insurance commissioner before using them. Approval is presumed if rates are not denied within a specified number of days, in case there is a deemer clause. In the file and use system, rates become effective immediately upon filing. The insurance commissioner, however, may subsequently disapprove the rates. A use and file system provides that rates must be filed with the insurance commissioner within a specified period of time after their first use. In the flex rating system insurers may increase or decrease rates within a certain percentage range. Prior approval is required only if the rate

¹³² *Id.*

¹³³ *Id.*; Robert M. Ferm & Amy C. Palmerlee, *Recent Developments in Public Regulation of Insurance*, 35 TORT & INS. L. J. 603, 607 (2000).

¹³⁴ Ferm & Palmerlee, *supra* note 133, at 607.

¹³⁵ *Id.*

¹³⁶ HARRINGTON, *supra* note 1, at 10; Ferm & Palmerlee, *supra* note 133, at 607.

¹³⁷ Ferm & Palmerlee, *supra* note 133, at 607; *see also* HARRINGTON, *supra* note 1, at 10 (stating that the deregulation trend reflected recognition that rate and form regulation serve no useful purpose and are instead counterproductive).

¹³⁸ NAIC, 2 COMPENDIUM OF STATE LAWS ON INSURANCE TOPICS, HEALTH/LIFE/PROPERTY/CASUALTY II-PA-10-21 (2011).

¹³⁹ *Id.*

change is larger than the specified percentage. The modified prior approval system provides that rate revisions based only on a change in loss experience are subject to file and use regulation. However, rate revisions based on a change in expense ratio or rate classifications are subject to prior approval. Lastly, under the no-filing system rates do not need to be filed with or approved by the state insurance commissioner. The table in Appendix 1 identifies the types of rate regulation systems according to state.

Recent changes in state rating laws confirm the trend towards less restrictive systems of rate regulation. As of April 2008 rates in Massachusetts were determined according to what the insurance commissioner at the time, Nonnie Burns, called “managed competition”.¹⁴⁰ Previously Massachusetts had been the only state where the insurance commissioner set rates for auto insurance. Now, insurance companies submit their rates to the state insurance commissioner, who has power to disapprove them if they are excessive or unfairly discriminatory.¹⁴¹ In May 2008 the Georgia legislature passed legislation¹⁴² permitting auto insurance companies to set rates above the mandatory minimum limits without prior approval from the insurance commissioner.¹⁴³ Finally, in June 2008 the New York legislature approved an auto insurance flex rating bill¹⁴⁴ that allows auto insurance companies to adjust their rates twice annually within a 5 percent band without seeking prior regulatory approval.¹⁴⁵

¹⁴⁰ Robert P. Spellane, *Managed Competition Good News for Auto Insurance Buyers*, <http://www.thefreelibrary.com/Managed+competition+good+news+for+auto+insurance+buyers.-a0168869015>.

¹⁴¹ Press Release, Mass. Gov’t, Office of Consumer Affairs & Bus. Regulation, Mass. Gov’t, Insurance Commissioner Issues Final Managed Competition Regulation (Oct. 5, 2007), http://www.mass.gov/?pageID=ocapressrelease&L=1&L0=Home&sid=Eoca&b=pressrelease&f=07_10_05_autoinsurance&csid=Eoca.

¹⁴² S.B. 276, 2008 Gen. Assemb. Sess. (Ga. 2009) (codified as amended at GA. CODE ANN. § 33-9-21(b)(21) (West 2011)).

¹⁴³ GA. CODE ANN. § 33-9-21(b)(2) (West 2011).

¹⁴⁴ N.Y. INS. LAW § 2350 (McKinney 2011).

¹⁴⁵ *Id.*

II. RATIONALES FOR AND AGAINST RATE REGULATION

A. PURPOSES OF RATE REGULATION

In the 1914 landmark case, *German Alliance Ins. Co. v. Lewis*, the U.S. Supreme Court upheld the power of the states to regulate rates on grounds that insurance is affected by the public interest.¹⁴⁶ According to the public interest theory and normative economic theory, insurance rate regulation is intended to correct market failures that would otherwise cause inefficiency and inequity in the insurance market and harm the interest of the general public.¹⁴⁷

In particular, rate regulation primarily aims to remedy two opposite problems: one, the tendency of insurance companies to engage in destructive competition and two, the formation of price cartels by insurance companies that could set excessive rates.¹⁴⁸ These aims were clearly defined by the subcommittee appointed by the NAIC Rates and Rating Organization Committee in 1960 to review the status of insurance rate regulation.¹⁴⁹ In a report dated November 28, 1960, the subcommittee stated that rate regulation is intended to assure that insurance coverages desired by the public are offered to the public by licensed insurers, that the cost of these coverages is reasonable and not excessive, that insureds bear a fair share of the cost of insurance and that insurers remain solvent to protect their policyholders.¹⁵⁰ In accordance with those objectives, states seek to promote the public welfare by ensuring that premiums are “adequate, not excessive and not unfairly discriminatory.” Although these three rate standards are interpreted differently in the different states, they have basic common features that the following analysis will describe.

The first aim of insurance rate regulation, to ensure adequate rates, stems from past experience with unregulated rates and the consequent destructive competition that led to several insurers’ insolvencies.¹⁵¹ One of the principal aims of rate regulation is to maintain the solvency of

¹⁴⁶ *German Alliance Ins. Co. v. Lewis*, 233 U.S. 389, 414-18 (1914).

¹⁴⁷ David J. Cummins, *Property-Liability Insurance Price Deregulation: The Last Bastion?*, in DEREGULATING PROPERTY-LIABILITY INSURANCE: RESTORING COMPETITION AND INCREASING MARKET EFFICIENCY 6 (Cummins ed., 2002).

¹⁴⁸ *Id.*

¹⁴⁹ See *supra* p. 122.

¹⁵⁰ HANSON ET AL., *supra* note 2, at 46-47.

¹⁵¹ See *supra* pp. 113-16.

insurance companies and to prevent rates from being set too low.¹⁵² The importance of this goal results from the unique nature of the insurance contract. An insurance contract, indeed, is an aleatory contract in which the policyholder pays up-front premiums in exchange for the insurer's promise to indemnify in case a future loss occurs. This gives policyholders an interest in the financial solidity of their insurance companies and in the companies' ability to pay future claims. Given the function of insurance in satisfying society's need for security, the financial solvency of insurance companies is the principal purpose of insurance regulation.¹⁵³ For this reason, "the principle of *solidity* is pervasive" in insurance regulation.¹⁵⁴ Rate regulation, like capital adequacy, is considered a means to ensure the solidity of the insurance industry.¹⁵⁵ The self-interest of insurance companies in remaining solvent has not always been reckoned sufficient.¹⁵⁶ Rate regulation, instead, is believed to assure insurers' solvency by keeping rates above a certain minimum level of adequacy so that adequate reserves can be maintained.¹⁵⁷

As for the second purpose of insurance rate regulation, the concern for making rates not excessive was not one of the original reasons for regulating prices. Initially, regulators and insurance companies were concerned about ruinous competition that could threaten solvency.¹⁵⁸ The NAIC-AIC model laws introduced the "not excessive" standard due to the drafters' belief that cooperation among insurers in setting rates created a need to prevent excessive rates.¹⁵⁹

The "not excessive" standard seeks to make the cost of insurance affordable.¹⁶⁰ In the process, the standard promotes the availability of insurance.¹⁶¹ At the same time, this standard may be in tension with the first

¹⁵² BANKS MCDOWELL, *DEREGULATION AND COMPETITION IN THE INSURANCE INDUSTRY* 36 (Greenwood Press 1989); Irwin M. Stelzer & Geraldine Alpert, *Benefits and Costs of Insurance Deregulation*, in *INSURANCE DEREGULATION: ISSUES AND PERSPECTIVES* 6, 8 (Nathan Weber ed., The Conference Board 1982).

¹⁵³ Spencer L. Kimball, *The Purpose of Insurance Regulation: A Preliminary Inquiry in the Theory of Insurance Law*, 45 MINN. L. REV. 471, 478-80 (1961).

¹⁵⁴ *Id.* at 480.

¹⁵⁵ *Id.* at 481-83.

¹⁵⁶ *Id.* at 482-83.

¹⁵⁷ MCDOWELL, *supra* note 152, at 36.

¹⁵⁸ *Id.* at 37.

¹⁵⁹ *Id.*

¹⁶⁰ Kimball, *supra* note 153, at 491.

¹⁶¹ Stelzer & Alpert, *supra* note 152, at 8.

standard, which requires insurance rates to be adequate in order to assure the solvency of the insurance companies.¹⁶²

Finally, rate insurance regulation seeks to avoid rates that unfairly discriminate among insureds. This third goal reflects regulators' concerns about price discrimination among consumers when such discrimination is not related to differences in the risks underwritten.¹⁶³ Objectionable forms of discrimination include: (i) unfair individual discrimination, such as rebates, credits and misclassifications that favor one insured over another when the risk underwritten is the same,¹⁶⁴ (ii) unfair group rate discrimination that usually involves rating plans that arbitrarily differentiate among the insureds without taking into account their risk¹⁶⁵ and (iii) unfair product discrimination that results in unreasonable overpricing or under pricing of one product compared to another.¹⁶⁶ In this regard, insurance regulators aim to ensure that rates are fair for every class of insured and that the classes are fair and nondiscriminatory.¹⁶⁷ Therefore, while the standard of "not excessive" rates seeks to accomplish reasonableness between insurance companies and policyholders, the standard of "not unfairly discriminatory" rates seeks to accomplish "equity" by ensuring that policyholders are not unfairly discriminated against.¹⁶⁸ In order to achieve this objective, fair classifications of policyholders for premium calculation are necessary so that every insured will bear the cost of his or her own insurance.¹⁶⁹ It is difficult to make fair classifications, however, since every risk is unique and theoretically could be uniquely rated.¹⁷⁰

¹⁶² *Id.*

¹⁶³ MCDOWELL, *supra* note 152, at 39; SHARON TENNYSON, EFFICIENCY CONSEQUENCES OF RATE REGULATION IN INSURANCE MARKETS, Networks Financial Institute 6 (2007), http://www.networksfinancialinstitute.org/Lists/Publication%20Library/Attachments/15/2007-PB-03_Tennyson.pdf.

¹⁶⁴ HANSON ET AL., *supra* note 2, at 432-33.

¹⁶⁵ *Id.* at 433.

¹⁶⁶ *Id.*

¹⁶⁷ MCDOWELL, *supra* note 152, at 39.

¹⁶⁸ Kimball, *supra* note 153, at 495.

¹⁶⁹ HANSON ET AL., *supra* note 2, at 93; Kimball, *supra* note 153, at 495; Jason C. Blackford, *Competition as a Means of Regulating Insurance*, 18 CLEV.-MARSHALL L. REV. 116, 128 (1969).

¹⁷⁰ Kimball, *supra* note 153, at 495.

B. ARGUMENTS FOR RATE REGULATION

Having analyzed the goals of rate regulation, now the article turns to the arguments in favor of state regulation of insurance rates.

According to the traditional rationale for regulation of insurance rates, states can protect policyholders by controlling rates.¹⁷¹ The argument is based on the fact that policyholders and insurance companies do not deal at arm's length and that insurance companies are likely to overcharge policyholders in the absence of rate regulation.¹⁷² When competition results in a variety of rates, some argue that policyholders do not benefit from that variety because they may not have the ability to compare the rates.¹⁷³ Policyholders have difficulty in fully understanding the insurance contract and in establishing a connection between the price and the quality of the coverage. In these circumstances, rate regulation and standardization are said to be appropriate.¹⁷⁴ In an un-regulated system, some insurers may cut rates by providing low-quality insurance products that policyholders might not recognize as poor quality due to imperfect information about the characteristics of the coverage offered and the financial solidity of the insurer.¹⁷⁵ Rate regulation would counteract deception by insurers and assist consumers in comparing different insurance policies.¹⁷⁶

Rate regulation also helps to prevent ruinous price competition with a subsequent increase in insurers' insolvencies.¹⁷⁷ Advocates for rate regulation argue that insurance companies will respond to the danger of destructive competition, by conspiring to set rates.¹⁷⁸ This raises concerns

¹⁷¹ JONATHAN R. MACEY & GEOFFREY P. MILLER, *COSTLY POLICIES: STATE REGULATION AND ANTITRUST EXEMPTION IN INSURANCE MARKETS* 106 (Christopher C. DeMuth & Jonathan R. Macey eds., American Enterprise Institute Press 1993).

¹⁷² *Id.*

¹⁷³ FREDERICK G. CRANE, *AUTOMOBILE INSURANCE RATE REGULATION* 84-85 (Ohio State University 1962); ROBERT E. DINEEN ET AL., *INSURANCE AND GOVERNMENT: THE ECONOMICS AND PRINCIPLE OF INSURANCE SUPERVISION* 14 (Charles C. Center & Richard M. Heins eds., University of Wisconsin 1960); HANSON ET AL., *supra* note 2, at 529-30; MACEY & MILLER, *supra* note 187, at 106.

¹⁷⁴ HANSON ET AL., *supra* note 2, at 529-30.

¹⁷⁵ *Id.*

¹⁷⁶ *Id.* at 530.

¹⁷⁷ Blackford, *supra* note 169, at 129.

¹⁷⁸ *Id.* at 130.

about anti-competitive conduct and the converse danger of excessive rates.¹⁷⁹

Predatory pricing concerns are another reason to favor of rate regulation. The concern is that rate deregulation might induce stronger insurers to use their greater financial resources to temporarily cut rates to increase their market share and force weaker insurers out of the insurance market.¹⁸⁰

Rate regulation is also urged in the interest of actuarial accuracy, because regulators must rely on wide loss experience in setting rates that even larger insurance companies may lack.¹⁸¹ Moreover, unregulated rates might lead to underwriting restrictions because some insurers might decide to write only low-risk insureds in order to minimize their costs and charge lower premiums.¹⁸² Insurers that continued to write higher-risk insureds would bear a greater proportion of such risks and be forced to increase their rates in order to cover possible losses.¹⁸³ This would create problems of insurance affordability and has the potential to result in insolvency of the higher-risk insurers.¹⁸⁴

C. ARGUMENTS AGAINST RATE REGULATION

This article will now examine the arguments in favor of deregulating rates. The principal rationale for insurance rate regulation is that it is needed to correct market failures.¹⁸⁵ Opponents of rate regulation, however, argue that the insurance market is competitive¹⁸⁶ since it is characterized by a large number of firms doing business with a low level of concentration and selling similar products.¹⁸⁷ They agree that there are modest barriers to entry and that profits are not excessive compared with

¹⁷⁹ *Id.*

¹⁸⁰ HANSON ET AL., *supra* note 2, at 528; Blackford, *supra* note 169, at 130.

¹⁸¹ HANSON ET AL., *supra* note 2, at 528.

¹⁸² CRANE, *supra* note 173, at 78-79; HANSON ET AL., *supra* note 2, at 529.

¹⁸³ CRANE, *supra* note 173, at 78-79; HANSON ET AL., *supra* note 2, at 529.

¹⁸⁴ CRANE, *supra* note 173, at 78-79; HANSON ET AL., *supra* note 2, at 529.

¹⁸⁵ See *supra* p. 127; HARRINGTON, *supra* note 1, at 15; Cummins, *supra* note 147, at 6; Tennyson, *supra* note 163, at 8; D'Arcy, *supra* note 110, at 262.

¹⁸⁶ See *infra* pp. 136-41, 150-52.

¹⁸⁷ HARRINGTON, *supra* note 1, at 15-18; Cummins, *supra* note 147, at 7; Tennyson, *supra* note 163, at 8; Paul L. Joskow, *Cartels, Competition and Regulation in the Property-Liability Insurance Industry*, 4 BELL J. OF ECON. AND MGMT. SCI. 375, 379-82, 391 (1973); Joskow & McLaughlin, *supra* note 108, at 378-79.

other industries.¹⁸⁸ Thus, they conclude there is no evidence of market failure that may justify insurance rate regulation.¹⁸⁹

Further, with regard to the two opposite concerns that, in the absence of rate regulation, insurers would engage in destructive competition or form cartels that would lead to excessive rates,¹⁹⁰ proponents of rate deregulation note that these concerns are outdated.¹⁹¹

The concept of destructive competition dates back to the nineteenth century¹⁹² and is no longer well founded since more recently there has been no evidence of dangerous price cutting; rather prices in insurance markets reflect expected claim costs and reasonable profits for insurance companies.¹⁹³ Those who favor deregulated rates stress that insurance companies, like all other enterprises, aim to conduct a financially successful business and to avoid charging rates that are too low to cover their costs.¹⁹⁴ Under this view, rate deregulation is not likely to cause ruinous competition because, even assuming that a big insurer reduces its rates in order to eliminate possible competitors, in the long run it will have to raise its rates to cover its costs.¹⁹⁵ In that event, new competitors, attracted by the possibility of making profit, will enter the market.¹⁹⁶

The cartel pricing concern originated from the initial bureau rate-making activities and the regulatory restrictions on deviations from the bureau rates in the 1950s and early 1960s.¹⁹⁷ Effective cartel pricing is now unlikely given the large number of insurers, ease of entry into the market and the decreased influence by rate bureaus in setting rates.¹⁹⁸

¹⁸⁸ HARRINGTON, *supra* note 1, at 15-18; Cummins, *supra* note 147, at 7; Tennyson, *supra* note 163, at 8; Joskow, *supra* note 187, at 388-91; Joskow & McLaughlin, *supra* note 108, at 379.

¹⁸⁹ HARRINGTON, *supra* note 1, at 18; Cummins, *supra* note 147, at 7; Tennyson, *supra* note 163, at 8; Joskow, *supra* note 187, at 403; Joskow & McLaughlin, *supra* note 108, at 379.

¹⁹⁰ *See supra* pp. 111-23.

¹⁹¹ Cummins, *supra* note 147, at 6.

¹⁹² *Id.*

¹⁹³ CRANE, *supra* note 173, at 87-88; Cummins, *supra* note 147, at 6.

¹⁹⁴ CRANE, *supra* note 173, at 88.

¹⁹⁵ *Id.*

¹⁹⁶ *Id.*

¹⁹⁷ Joskow & McLaughlin, *supra* note 108, at 379-80; Cummins, *supra* note 147, at 7.

¹⁹⁸ Joskow & McLaughlin, *supra* note 108, at 379-80; Cummins, *supra* note 147, at 7; HANSON ET AL., *supra* note 2, at 447-53.

Advocates of rate deregulation maintain that the free market is just as appropriate for the insurance sector as it is for other businesses.¹⁹⁹ Although insurance regulation is considered important to ensure the protection of policyholders, proponents of deregulation agree that state control of the insurance market impedes competition.²⁰⁰ In this connection, rate regulation may lead to both inadequate and excessive rates.²⁰¹ In the latter case, rate regulation might cause levels of premiums so high that even the most inefficient insurance companies would make profits.²⁰² Regulators should not intervene in rate setting and insurance companies should be permitted to make rates so that policyholders can benefit from lower-cost insurance.²⁰³

Further, advocates point out that determining a proper rate is not feasible since rate setting is “not an inevitably accurate and scientific calculation.”²⁰⁴ They observe that rate-setting is a subjective activity and because there can be more than one reasonable decision in making rates, there is no reason to regard a commissioner’s decision as the most reasonable.²⁰⁵ Indeed, the setting of rates by competently managed insurance companies is arguably as reasonable as the setting of rates by the commissioners.²⁰⁶ It is also emphasized that a rate proper for one insurer might not be proper for another one.²⁰⁷

In addition, proponents argue that unregulated rates will avoid commission wars.²⁰⁸ When price uniformity prevails, insurers are more likely to have to pay agents higher commissions in order to obtain business.²⁰⁹ Conversely, they say, the problem of commission wars can be overcome when rates are deregulated since insurers can obtain business by competing on the price of the products offered.²¹⁰ Another disadvantage is that restrictions on price competition limit product differentiation because

¹⁹⁹ CRANE, *supra* note 173, at 84; HANSON ET AL., *supra* note 2, at 530.

²⁰⁰ CRANE, *supra* note 173, at 84; HANSON ET AL., *supra* note 2, at 530.

²⁰¹ HANSON ET AL., *supra* note 2, at 535-36.

²⁰² HANSON ET AL., *supra* note 2, at 530.

²⁰³ *Id.*

²⁰⁴ CRANE, *supra* note 173, at 86 (citing Thomas O. Carlson, President, “Statistics for the Ratemaker,” *Proceedings of the Casualty Actuarial Society* 1 (May 1953)).

²⁰⁵ *Id.*

²⁰⁶ *Id.*

²⁰⁷ *Id.*

²⁰⁸ HANSON ET AL., *supra* note 2, at 532.

²⁰⁹ *Id.*

²¹⁰ *Id.*

comparable rates must be charged for comparable products in order to implement a uniform pricing system.²¹¹ Moreover, rate regulation requires even more resources and efforts by the insurance departments in order to examine insurance rates. Deregulation of rates would permit regulators to fully devote themselves to other more important supervisory activities such as solvency supervision.²¹²

Another reason advanced in favor of rate deregulation is that it would take politics out of rate-setting.²¹³ Rate regulation often involves political pressure on insurance commissioners by insurers demanding rate increases and consumers that look unfavorably on those increases.²¹⁴ In particular, advocates of rate deregulation observe that the political pressure by policyholders may lead to inadequate rates since regulators will be influenced to approve rate increases that “may be either too little and/or too late.”²¹⁵ Ironically, even though rate regulation is aimed at avoiding inadequate rates, it may actually lead to inadequate rates.²¹⁶ This is especially true in prior-approval systems, because of delays in obtaining approval cost insurers, especially after taking inflation into account.²¹⁷ Insurance companies react to inadequate rates by restricting underwriting or by cancelling and refusing to renew insurance policies creating subsequent possible problems of unavailable coverage.²¹⁸ This result undermines one of the purposes of insurance rate regulation, to promote insurance availability. Consequently, reduced rate regulation will give

²¹¹ CRANE, *supra* note 173, at 96; HANSON ET AL., *supra* note 2, at 533.

²¹² See CRANE, *supra* note 173, at 97-99; HANSON ET AL., *supra* note 2, at 533; HARRINGTON, *supra* note 1, at 33.

²¹³ CRANE, *supra* note 173, at 100; HANSON ET AL., *supra* note 2, at 62-64, 535.

²¹⁴ CRANE, *supra* note 173, at 99-100; HANSON ET AL., *supra* note 2, at 62-64, 535.

²¹⁵ HANSON ET AL., *supra* note 2, at 64; David J. Cummins, Richard D. Phillips & Sharon Tennyson, *Regulation, Political Influence and the Price of Automobile Insurance*, 20 J. INS. REG. 9, 42-44 (2001) (showing by statistical regression analysis that insurance prices in regulated states are affected by political influence activities of consumer groups); David J. Cummins & Scott E. Harrington, *The Impact of Rate Regulation in U.S. Property-Liability Insurance Markets: A Cross-Sectional Analysis of Individual Firm Loss Ratios*, 12 GENEVA PAPERS ON RISK & INS. 50, 60 (1987) (suggesting that regulators responded to consumer pressure by holding rates below levels that would have occurred under pricing freedom); Harrington, *supra* note 118, at 189.

²¹⁶ HANSON ET AL., *supra* note 2, at 64-65, 535.

²¹⁷ *Id.* at 64.

²¹⁸ *Id.* at 535-36; Harrington, *supra* note 118, at 189.

insurers the flexibility to adjust rates, ensure adequate rates and make insurance available.²¹⁹

Finally, advocates of deregulation note that the prior approval system can cause rates to remain at a higher level than appropriate.²²⁰ Due to the time necessary to approve a new rate, a cost decline does not automatically translate into a lower rate.²²¹ Moreover, insurers may not apply for lower rates based on improvement in their underwriting experience if they expect to have difficulty in later obtaining a needed increase.²²²

III. THE EFFECTS OF OPEN COMPETITION

A. THE STRUCTURE OF THE U.S. PROPERTY AND CASUALTY INSURANCE MARKET

The economic justification for rate regulation is that it protects the public interest by avoiding inefficiency that would otherwise result from monopolistic or oligopolistic conduct. Under this logic, in order to assess whether the property and casualty industry is likely to achieve benefits from rate deregulation, market structure and easy of entry should be examined.²²³ The principal characteristic of monopoly is the presence of a single seller of a product for which there are no alternatives.²²⁴ However, economists have demonstrated that oligopoly power may also exist if there are few sellers and they act in concert.²²⁵ Entry by competitors is the main limitation on monopoly power in a market economy.²²⁶

The U.S. property and casualty insurance market has the structural characteristics of a competitive market.²²⁷ The market is characterized by a large number of firms operating with low levels of concentration and selling products with identical features.²²⁸ The competitive structure of the market is apparent from the fact that in 2009 there were 2,737 property and

²¹⁹ HANSON ET AL., *supra* note 2, at 536.

²²⁰ *Id.* at 71.

²²¹ *Id.*

²²² *Id.*

²²³ HARRINGTON, *supra* note 1, at 15.

²²⁴ KAHN, *supra* note 1, at 116.

²²⁵ *Id.*

²²⁶ *Id.*

²²⁷ HARRINGTON, *supra* note 1, at 16; Joskow, *supra* note 187, at 391.

²²⁸ Joskow, *supra* note 187, at 390.

casualty insurance companies operating in the United States.²²⁹ This number has increased since 1971, when there were 1,206 companies operating in the U.S. property and casualty insurance market.²³⁰

That said, the presence of a large number of insurers offering basically the same product is not by itself indicative of competition since a small number of companies could write a majority of the premiums and by virtue of their market share be able to fix prices.²³¹ It is necessary, therefore, to consider the relative market share of insurance companies in order to determine whether the market is competitive. The following table presents the 2009 nationwide market share of the top twenty-five U.S. property and casualty insurance groups. The market share of different corporate groups as a whole is a more accurate indicator than market share of their individual insurance subsidiaries. While individual subsidiaries are separate legal entities, they are not economically independent and are subject to the group's management decisions.

Table 1 – Property and Casualty Insurance Industry 2009 Market Share Nationwide by Group

	GROUP NAME	DIRECT PREMIUMS WRITTEN	MARKET SHARE percent	CUMULATIVE MARKET SHARE percent
1	State Farm Grp	51,063,110,761	10.50	10.50
2	Zurich Ins Grp	28,979,691,684	5.96	16.46
3	Allstate Ins Grp	26,153,440,231	5.38	21.84
4	American Intl Grp	26,140,201,178	5.38	27.22
5	Liberty Mut Grp	24,772,894,328	5.10	32.32
6	Travelers Grp	21,409,548,242	4.40	36.72
7	Berkshire Hathaway Grp	16,054,658,656	3.30	40.02
8	Nationwide Corp Grp	15,405,561,636	3.17	43.19
9	Progressive Grp	14,200,294,349	2.92	46.11
10	Hartford Fire & Cas Grp	10,473,026,375	2.15	48.26
11	United Serv Automobile Ass'n Grp	10,439,501,509	2.15	50.41

²²⁹ INSURANCE INFORMATION INSTITUTE, THE INSURANCE FACT BOOK 2011 V (2011) [hereinafter THE INSURANCE FACT BOOK 2011].

²³⁰ Joskow, *supra* note 187, at 379.

²³¹ *Id.* at 380.

12	Chubb & Son Ins Grp	9,419,255,363	1.94	52.35
13	Cna Ins Grp	8,131,205,861	1.67	54.02
14	Ace Ltd Grp	7,780,534,083	1.60	55.62
15	Allianz Ins Grp	5,764,589,841	1.19	56.81
16	American Family Ins Grp	5,681,564,588	1.17	57.98
17	Auto Owners Grp	4,451,729,312	0.92	58.90
18	Erie Ins Grp	3,860,839,234	0.79	59.69
19	Assurant Inc Grp	3,735,278,486	0.77	60.46
20	American Financial Grp	3,565,868,308	0.73	61.19
21	Wr Berkley Corp Grp	3,255,838,299	0.67	61.86
22	Fm Global Grp	3,199,857,312	0.66	62.52
23	Qbe Ins Grp	3,128,630,118	0.64	63.16
24	Cincinnati Fin Grp	3,071,344,125	0.63	63.79
25	Metropolitan Grp	2,984,332,558	0.61	64.40

Source: NAIC, 2009 Market Share Reports for the Top 25 Property/Casualty Insurers Over 25 Years 39 (2010), reprinted with permission. The NAIC does not endorse any analysis or conclusions based on use of its data.

Although table 1 suggests that the market is concentrated since twenty-five insurance groups control 64.40 percent of the market, with State Farm Group controlling a market share of 10.50 percent,²³² evaluation of the Herfindahl-Hirschman Index (HHI) leads to a different conclusion. The HHI is a commonly used measure of industry concentration and is calculated by summing the squares of the market share percentage of all companies in the market. For example, if a market had only one seller, its market share would be 100 percent and its HHI would be 10,000. If a market had five sellers, each with an equal 20 percent of the market, the HHI would be 2000. The HHI tends to zero when a market consists of a large number of firms of relatively equal size. Increases in the value of the HHI indicate higher concentration in the market, either due to a decrease in the number of firms or an increase in the disparity in size between these firms. Although there is no precise point at which the HHI indicates market concentration sufficient to restrict competition, the Department of Justice

²³² See James Barrese, Gene Lai & Nicos Scordis, *Ownership Concentration and Governance in the U.S. Insurance Industry*, 30 J. INS. ISSUES 1 (2007).

has developed Merger Guidelines under which an HHI of less than 1000 means the market is not concentrated, an HHI between 1,000 and 1,800 points means the market is moderately concentrated and an HHI over 1,800 points means the market is concentrated.²³³ According to the ISO, the HHI for the property and casualty insurance market in 2009 was 351 points.²³⁴ This indicates that the market was not concentrated. Further, the trend toward an increase in the level of market concentration indicates a decline in high-cost companies in favor of more efficient and lower-cost companies.²³⁵ Therefore, higher market concentration may be the result of increased market competition and a subsequent improvement in policyholders' welfare.²³⁶

With respect to possible barriers to entry, it is generally acknowledged that insurers can easily enter the property and casualty insurance market.²³⁷ The ability of new insurers to enter into the business assures efficiency and competition. When there are excessive profits in the market, new firms are induced to enter and the quantity of products offered is increased. Consequently, excess profits decrease until reaching a price level where zero excess profits exist. In this way a competitive market is achieved. The following table shows, *inter alia*, the number of entries in the markets for commercial property and casualty insurance products from 2004 to 2009.

Table 2 – 2009 Commercial Lines Data – Nationwide

LINE OF BUSINESS	PREMIUMS WRITTEN	MARKET SHARE FOUR LARGE GROUPS	HHI BASED ON PREMIUM	NUMBER OF SELLERS (GROUPS)	NUMBER OF ENTRIES LAST 5 YEARS	NUMBER OF EXITS LAST 5 YEARS	MARKET GROWTH LAST 3 YEARS	MARKET GROWTH LAST 10 YEARS	RETURN ON NET WORTH 10-YEAR MEAN
Commercial Auto Liability	18,988,326,402	28.55 percent	320	105	30	29	-13.97 percent	25.05 percent	7.43 percent

²³³ U.S. Department of Justice and the Federal Trade Commission, Horizontal Merger Guidelines § 1.51 (1997), *available at* <http://www.usdoj.gov/atr/public/guidelines/hmg.htm>.

²³⁴ THE INSURANCE FACT BOOK 2011, *supra* note 229, at 47.

²³⁵ Joskow, *supra* note 187, at 382.

²³⁶ *Id.*

²³⁷ HARRINGTON, *supra* note 1, at 16; Joskow, *supra* note 187, at 388-91; Joskow & McLaughlin, *supra* note 108, 379.

Commercial Auto Physical	5,792,918,385	25.21 percent	266	119	32	24	-20.16 percent	-4.67 percent	13.36 percent
Commercial Auto Total	24,781,244,787	27.77 percent	297	115	33	25	-15.50 percent	16.56 percent	8.49 percent
Commercial Multiple Peril	34,034,902,544	27.80 percent	338	104	25	27	-5.82 percent	50.34 percent	8.03 percent
Fire	12,861,192,843	38.65 percent	554	99	33	33	5.11 percent	138.66 percent	19.77 percent
Allied Lines	11,249,248,316	38.47 percent	499	84	28	43	-1.31 percent	186.42 percent	-3.92 percent
Inland Marine	13,434,863,829	35.08 percent	495	76	25	32	-12.81 percent	61.49 percent	19.07 percent
Mortgage Guaranty	5,449,184,963	69.45 percent	1,594	8	1	1	-11.43 percent	46.31 percent	-33.39 percent*
Financial Guaranty	1,922,896,601	89.54 percent	2,985	9	5	5	-45.91 percent	22.40 percent	-15.44 percent*
Medical Professional Liability	10,817,257,976	24.28 percent	288	98	27	25	-7.42 percent	67.72 percent	7.40 percent
Other Liability	47,489,981,386	33.13 percent	451	88	23	21	-13.41 percent	82.38 percent	4.66 percent
Workers Compensation	41,287,350,051	33.34 percent	395	110	39	31	-20.98 percent	19.64 percent	6.35 percent
Products Liability	2,895,299,149	30.42 percent	404	66	19	16	-28.72 percent	47.40 percent	0.43 percent*

* Denotes Return on Net Worth for 2009 data year only.

Source: NAIC, 2009 Competition Database Report 10 (2010), reprinted with permission. The NAIC does not endorse any analysis or conclusions based on use of its data.

In particular, it can be seen that the number of insurance groups with affiliated insurers which have entered the markets for commercial property and casualty insurance products between 2004 and 2009 is substantial. For example, thirty-nine insurers entered the workers compensation market, thirty entered the commercial auto insurance market, twenty-seven entered the medical professional liability market and twenty-

five entered the commercial multiple peril market. This, along with the fact that the level of concentration in the U.S. property and casualty insurance market is low, leads to the conclusion that insurance companies are unable to charge excessive prices by attempting to act in concert since, in the absence of substantial barriers to entry, new insurers will prevent existing companies from fixing prices.

As this shows, the U.S. property and casualty insurance market is competitive because it is characterized by a large number of insurance companies operating with low concentration levels. Prof. Paul Joskow called the insurance market one of the markets that conform more closely to the ideal model of perfect competition.²³⁸ The insurance market, therefore, does not present characteristics of a monopoly or an oligopoly that may justify rate regulation.

In recognition of the wisdom of rate deregulation, there has been a gradual movement away from prior-approval systems toward less restrictive systems such as: file and use, use and file, flex rating, modified prior approval and, in particular, no file systems.²³⁹ Further, the NAIC File and Use Model Act introduced a presumption in favor of the existence of a competitive market unless the commissioner, after a hearing, determines that the market is not competitive.²⁴⁰ The Model Act also established a standard which provides that a rate in a competitive market is not excessive.²⁴¹

Nevertheless, prior-approval laws are still enforced in many states.²⁴² For example, prior approval systems are used in Mississippi with regard to all insurance lines, in California with regard to all lines except title insurance, in Alabama for medical malpractice, property and inland marine, workers' compensation and personal lines, in Alaska with regard to medical malpractice, workers' compensation and assigned risk rates, in Connecticut with regard to medical malpractice (for rate increases of 7.5

²³⁸ Joskow, *supra* note 187, at 391.

²³⁹ *See supra* pp. 125-26.

²⁴⁰ *E.g.*, PROPERTY AND CASUALTY MODEL RATING LAW (FILE AND USE VERSION) § 4 (NAIC 2010) [hereinafter FILE AND USE MODEL LAW] (providing that the insurance commissioner in determining whether a reasonable degree of competition exists in the market shall consider market structure, market performance, market conduct, the consumers' practical opportunities to acquire pricing and other information and to compare and purchase insurance from competing insurers).

²⁴¹ *See, e.g.*, FILE AND USE MODEL LAW, *supra* note 240, at § 5(A)(1)(a).

²⁴² *See* Appendix 1.

percent or more over the last rates filed) and title insurance.²⁴³ It is also worth mentioning that, except for the no-file systems, all the other systems mentioned above retain some form of regulatory control over insurance rates. Although rating laws in the different states vary to some extent, insurance commissioners retain the right to disapprove rates in file and use and use and file systems, while in flex rating and in modified prior approval systems insurers may be required to obtain prior approval from a commissioner if an increase is larger than the percentage rate established or the rate revision is based on a change in expense ratio or rate classifications.²⁴⁴ With a few rare exceptions, purely competitive rating models are not used in the United States. No-file systems are limited to just a few lines in some states.²⁴⁵ This stands in contrast with the fact that the competitive structure of the property and casualty insurance market in the U.S. does not justify the regulation of rates.

B. THE EUROPEAN EXPERIENCE WITH REGARD TO REGULATION OF INSURANCE TARIFFS

European Member States' regulators do not have the right to regulate insurance prices. It is worthwhile, therefore, to analyze the EU experience in order to draw possible conclusions that could be valuable in considering rate deregulation.

Insurance regulation in Europe aims to create an integrated insurance market so that insurers can better diversify their risks and attain more economies of scale, while allowing policyholders to benefit from increased competition and a wider choice of insurance products.

To this end, the EU legislature has attempted to remove regulatory barriers between Member States by introducing the principles of freedom of establishment and freedom to provide services.²⁴⁶ In order to foster

²⁴³ See Appendix 1.

²⁴⁴ See *supra* p. 126.

²⁴⁵ See Appendix 1.

²⁴⁶ The third non-life Council Directive 92/49/EEC, 1992 O.J. (L 228) and the third life Council Directive 92/96/EEC, 1992 O.J. (L 360) established a single system for the authorization and financial supervision of insurance companies by the Member State in which an insurer has its head office (the home Member State). The authorization issued by the home Member State allows an insurance company to conduct its business in the other European Member States, either by opening agencies or branches (freedom of establishment) or by offering services on a temporary basis (freedom to provide services). In general, the principle of freedom

competition in the single insurance market, the third non-life insurance Directive prevented insurance supervisory authorities from regulating insurance premium prices and policy conditions.²⁴⁷

Previously, most EU Member States had exercised considerable control over premiums by setting minimum or maximum prices or fixing price scales for some lines of insurance or even for all insurance lines.²⁴⁸ In the Italian insurance market, for example, before the enactment of the “third generation” of Directives the principles of “authorization of admission” and of “control on tariffs” were well established in the industry.²⁴⁹ Before deregulation, potential competition in the European insurance market was impeded by regulated tariffs that hampered insurance companies from competing on price.²⁵⁰

With the removal of national control over insurance tariffs, new insurance products can be introduced into the market without prior regulatory approval. In this way, insurers’ efficiency increased and consumers benefited from lower prices.²⁵¹ Article 29 of the third non-life insurance Directive of 1992 provides that Member States cannot maintain or introduce systems of prior notification or approval of insurers’ proposed increases in premium rates except as a part of general systems aimed at controlling prices.²⁵² Insurance companies in Europe are now free to set their rates without any state interference and to write insurance contracts on any terms they agree to with their policyholders. Efforts by Member States to control insurance prices have been censured by the European Commission. In 2000 the Italian government, due to the effects of motor insurance prices on inflation, imposed a one-year ban on any increase in premiums for certain policyholders whose rates were calculated on the

of establishment and freedom to provide services are set out, respectively, in article 49 and article 56 of the Treaty on the Functioning of the European Union.

²⁴⁷ Council Directive 92/49/EEC, art. 29, 1992 O.J. (L 228).

²⁴⁸ European Commission, *Business Insurance Sector Inquiry* 20 (2007), http://ec.europa.eu/competition/sectors/financial_services/inquiries/interim_report_24012007.pdf.

²⁴⁹ Giuseppe Turchetti & Cinzia Daraio, *How Deregulation Shapes Market Structure and Industry Efficiency: The Case of the Italian Motor Insurance Industry*, 29 GENEVA PAPERS ON RISK AND INS. 202 (2004).

²⁵⁰ European Commission, *supra* note 248, at 36.

²⁵¹ *Id.* at 45.

²⁵² *See also* Council Directive 92/49/EEC, art. 29, 1992 O.J. (L 228) 16-17.

basis of accidents.²⁵³ Additionally, the Italian government imposed a one-year freezing on all new policies that were calculated on the same basis.²⁵⁴ The European Commission censured the measure.²⁵⁵ According to the Commission, the price freeze was incompatible with the freedom to market insurance products within the European Union under the third non-life insurance Directive and was neither part of a general price-control system nor was it justified by the public interest.²⁵⁶

There is a legitimate concern that deregulation will obstruct setting accurate rates in the short run because insurers may not have sufficient loss experience on which to rely.²⁵⁷ The European Union addressed this problem within the framework of the insurance Block Exemption Regulations. The first Block Exemption Regulation, Regulation 3932/92, was adopted by the Commission in 1992.²⁵⁸ When this Regulation expired on March 31, 2003, the Commission replaced it with Regulation 358/2003.²⁵⁹ Afterwards, when also this second Regulation expired, on March 31, 2010, the Commission adopted a new insurance block exemption Regulation, Regulation 267/2010.²⁶⁰ The first Block Exemption Regulation was introduced following the *Verband der Sachversicherer* case, in which the European Court of Justice rejected arguments that full competition would cause more insurers' insolvencies and that, since cooperation between insurance companies was necessary to avoid such a risk, the applicability of Article 101 of the Treaty on the Functioning of the European Union (formerly Article 81 of the Treaty establishing the European Community)²⁶¹ should be limited.²⁶² Article 101(1) of the Treaty

²⁵³ If no accidents caused by the policyholders had occurred during a recent observation period. *See* art. 2, Legge 26 maggio 2000, n. 137, in G.U. 27 maggio 2000, n. 122 (It.).

²⁵⁴ *Id.*

²⁵⁵ *See* Press Release, Motor Insurance: The Commission Decides to Ask Italy to End Its Price Freeze (Oct. 25, 2000), <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/00/1210&format=HTML&aged=1&language=EN&guiLanguage=en>.

²⁵⁶ *Id.*

²⁵⁷ *See supra* p. 132.

²⁵⁸ *See* Commission Regulation 3932/92, art. 1, 1992 O.J. (L 398) 9 (EC).

²⁵⁹ *See* Commission Regulation 358/2003, art. 1, 2003 O.J. (L 53) 11-12 (EC).

²⁶⁰ *See* Commission Regulation 267/2010 2010, art. 2, 2010 O.J. (L 83) 5 (EU). This Regulation will expire on 31 March 2017. *See* Commission Regulation 267/2010 2010, art. 9, 2010 O.J. (L 83) 7 (EU).

²⁶¹ With effect from 1 December 2009, Article 81 of the EC Treaty has become Article 101 of the Treaty on the Functioning of the European Union. The

prohibits, *inter alia*, agreements between undertakings that prevent, restrict or distort competition within the EU common market by fixing prices and other trading conditions either directly or indirectly prices.²⁶³

The Commission, recognizing the importance of cooperation among insurance companies to produce pool data concerning the calculation of the average cost of covering a specified risk in the past, the frequency and the size of past insurance claims, exempted the joint compilation and distribution of calculations and studies from the application of article 101(1) of the Treaty.²⁶⁴ The Commission also exempted other agreements in the insurance sector concerning the setting up and operation of industry (re)insurance pools for the common coverage of certain risks in the form of co-(re)insurance.²⁶⁵

two articles are substantially the same. See Commission Regulation 267/2010, art. 1, fn. 2, 2010 O.J. (L 83) 1 (EU). Article 101(1) prohibits “all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition.” See Consolidated Version of the Treaty on the Functioning of the European Union art. 101(1), Mar. 30, 2010, 2010 O.J. (C 83), 88. As an exception to this rule, Article 101(3) provides that the provisions contained in Article 101(1) may be declared inapplicable in case of agreements “which contribut[e] to improving the production or distribution of goods or to promoting technical or economic progress, while allowing consumers a fair share of the resulting benefit[s]”, and which do not impose restrictions which are not indispensable to the attainment of these objectives and do not afford such undertakings the possibility of eliminating competition in respect of a substantial part of the products concerned. See Consolidated Version of the Treaty on the Functioning of the European Union art. 101(3), Mar. 30, 2010, 2010 O.J. (C 83), 88-89.

²⁶² See Case 45/85, *Verband der Sachversicherer e.V. v. Comm’n of the European Communities*, 1987 E.C.R. 405.

²⁶³ See Consolidated Version of the Treaty on the Functioning of the European Union art. 101(1), Mar. 30, 2010, 2010 O.J. (C 83) 88.

²⁶⁴ See Commission Regulation 267/2010, art. 2, 2010 O.J. (L 83) 5 (EU). According to paragraph 9 of the Preamble, access to past statistical data is essential in order to facilitate the pricing of risks and therefore the Commission considered cooperation in this area necessary. This can in turn facilitate market entry and benefit consumers. It is specified, however, that agreements on commercial premiums are not exempted. See also Commission Regulation 267/2010, pmb. ¶ 9, 2010 O.J. (L 83) 2 (EU).

²⁶⁵ See Commission Regulation 267/2010, art. 5, 2010 O.J. (L 83) 6 (EU). Commission Regulation 267/2010 did not renew the exemption granted by the previous Block Exemption Regulation for agreements on standard policy conditions and security devices. In particular, according to paragraph 3 of the

Unlike the old system where regulated tariffs prevented price competition, liberalization following the third non-life insurance Directive led to increased competition, particularly in formerly heavily-regulated markets such as Italy, Germany, Belgium and Portugal.²⁶⁶ After price controls were abolished, premium rates decreased.²⁶⁷ In particular, in countries such as Germany, Austria and Spain that used to have minimum premium regulation, price competition increased considerably.²⁶⁸ In Germany, due to discounts and price reductions, premium income from motor insurance decreased from DEM 44 billion in 1995 to DEM 39 billion in 1998.²⁶⁹ On the other hand, in countries such as Italy, Portugal and Greece, deregulation led to tariff increases in the motor liability sector in order to cover actual claim costs.²⁷⁰ Deregulation permitted insurance companies in those countries to reach a balance between the risks underwritten and the premiums charged to cover potential losses. Previously, in those countries premiums had been artificially low in order to prevent inflationary pressure.²⁷¹

The experience in countries like the United Kingdom and France, countries that had not regulated insurance prices and contractual terms before the Third Non-Life Insurance Directive confirms the benefits of rate deregulation.²⁷² In the United Kingdom, for example, market concentration has decreased as an immediate effect of deregulation. In 1981, the fifteen largest insurers operating in that country underwrote almost 80 percent of

Preamble, the new Regulation does not grant an exemption for the establishment of standard policy conditions and the testing and acceptance of security devices because the Commission's review of the functioning of Regulation 358/2003 revealed that it was no longer necessary to include such agreements in a sector specific block exemption regulation. The Commission considered more appropriate that they be subject to self-assessment. *See also* Commission Regulation 267/2010, pmbl. ¶ 3, 2010 O.J. (L 83) 1 (EU).

²⁶⁶ Autorità Garante della Concorrenza e del Mercato, *Gli Effetti Attesi della Liberalizzazione* 15-16 (2001), http://www.agcm.it/trasp-statistiche/doc_download/163-parte-i.html.

²⁶⁷ Swiss Re, Sigma: Europe in Focus: Non-Life Markets Undergoing Structure Change, 3 (2000).

²⁶⁸ *Id.*

²⁶⁹ *Id.*

²⁷⁰ *Id.*

²⁷¹ *Id.*

²⁷² Autorità Garante della Concorrenza e del Mercato, *Il Mercato*, 17 (2001), www.agcm.it/trasp-statistiche/doc_download/164-parte-ii.html (also stating that more recently the concentration ratios has increased due to insurance companies' reorganizations that have occurred recently).

the total premiums while in 1994, they underwrote about 65 percent of the total premiums.²⁷³ Moreover, as a consequence of market liberalization, even more foreign insurance companies set up business in the United Kingdom.²⁷⁴

Following the deregulation of insurance prices and conditions, concentration, however, did not decrease. After deregulation in 1992, the largest insurers consolidated their positions in their national markets. Between 1990 and 1998 the combined market share of European multinational insurers (Allianz, Axa, Cgu, Generali, Royal & Sun Alliance, Winterthur and Zurich) in the six largest national markets (United Kingdom, Germany, France, Italy, Netherlands, Spain) increased from 18 to 39 percent.²⁷⁵ In France in 1990, the combined market share of the five biggest insurance companies was about 40 percent, while in the second half of the nineties their market share increased to 57 percent.²⁷⁶ The same trend appeared in Italy. In 1990, the top five insurers controlled almost half of the Italian market, while in 1999 they had a market share of 60 percent.²⁷⁷ In Germany, the top five insurers had a market share of almost 32 percent in 1990 and 40 percent in 1999.²⁷⁸ In the United Kingdom, the top five insurers controlled 32 percent of the market in 1994 and 55 percent in 1999.²⁷⁹

It is difficult to know whether deregulation accounts for that higher concentration. More likely, the reduction in the number of insurers in the European market resulted from the increasing number of mergers and acquisitions at the end of the 1990s.²⁸⁰ For example, higher concentration ratios in Italy were due to the fact that Generali bought out INA in 1999, while in Germany they ensued from Generali's acquisition of AMB and AXA's takeover of Albingia.²⁸¹ Thus, it is unlikely that the increase in concentration ratios of the non-life European market resulted from deregulation. A case history of the motor insurance industry in Italy

²⁷³ *Id.*

²⁷⁴ *Id.*

²⁷⁵ Swiss Re, *supra* note 267, at 17.

²⁷⁶ *Id.* at 22.

²⁷⁷ *Id.* at 23-24.

²⁷⁸ *Id.*

²⁷⁹ *Id.* at 24.

²⁸⁰ See European Insurance in Figures, CEA STATISTICS N° 36 (CEA Insurers of Europe, Brussels, Belgium available at http://www.cea.eu/uploads/Modules/Publications/1224519688_eif.pdf), Oct. 2008, at 25 [hereinafter 2008 CEA Statistics].

²⁸¹ Swiss Re, *supra* note 267, at 23-24.

following deregulation is also illustrative. The Italian auto insurance sector, which had been highly regulated by the government, was considerably affected by the change introduced by the third non-life Directive.²⁸² The same trend seen in the general European non-life insurance market appeared in the Italian auto insurance market. Between 1982 and 1991, the number of insurance companies grew from 97 to 113, but then dropped to 80 between 1991 and 2000.²⁸³ The peak of the reduction occurred in the period after 1994, when the number of insurers declined from 105 to 80.²⁸⁴ Entries in the market rose in the second half of the 1980s and decreased in the 1990s.²⁸⁵ Conversely, the number of exits from the market decreased in the second half of the 1980s and increased in the 1990s.²⁸⁶ The net entry in the market between 1994 and 2002 was -28.²⁸⁷ The combined market share of the top 20 insurers also increased from 63.63 percent in 1982 to 79.87 percent in 2000.²⁸⁸ From this, one might infer that deregulation in the Italian auto insurance market had a negative effect on competition. However, in a study conducted in 2001, the Italian Antitrust Authority concluded that net exits from the market were not due to deregulation because only some of the insurance companies that exited the market had financial problems.²⁸⁹ Rather, the exits occurred because insurance companies were acquired by other companies and some insurers voluntarily ceased trading.²⁹⁰ The number of insurers' insolvencies decreased with the deregulation of insurance tariffs. In 1993-1994 around ten companies were insolvent, but in 1995 that number dropped to six.²⁹¹ The reduction in the number of insurers' insolvencies might be a result of the fact that insurers were free to set the price of premiums at an adequate level to cover their costs. Indeed, one adverse effect of rate regulation is to weaken the relationship between premiums and expected loss costs;²⁹² deregulation, on the contrary, permits a better alignment of prices with costs.²⁹³ Premiums rates went up and down until the first half of 1990s, while after tariffs

²⁸² See Turchetti & Daraio, *supra* note 249, at 202.

²⁸³ *Id.* at 203-04.

²⁸⁴ *Id.* at 204.

²⁸⁵ *Id.*

²⁸⁶ *Id.* at 205.

²⁸⁷ Autorità Garante della Concorrenza e del Mercato, *supra* note 272, at 35.

²⁸⁸ *Id.* at 26.

²⁸⁹ *Id.* at 36.

²⁹⁰ *Id.* at 36-37.

²⁹¹ *Id.* at 36.

²⁹² Cummins, *supra* note 147, at 12; see Tennyson, *supra* note 177, at 14.

²⁹³ See Cummins, *supra* note 147, at 2, 11.

deregulation they tended to increase.²⁹⁴ Among the causes adduced to explain rates increase are (1) the rise in the average cost of compensation for damage that changed from € 1,923 to € 3,830 between 1994 and 2001; (2) the increase in cost of repairs; (3) the frequency of fraud and the considerable frequency of cervical spine lesions reported in around 66 percent of the claims.²⁹⁵ Moreover, an efficiency analysis of forty-five Italian insurers in the motor insurance sector showed that the cost efficiency and the total productivity of these companies increased between 1982 and 2000, particularly in the second half of the 1990s after adoption of the third non-life Directive.²⁹⁶

Motor insurance aside, the other non-life lines in Italy experienced a decrease in rates from 1993 to 1996.²⁹⁷ This trend toward lower rates was common throughout Europe as a consequence of increased competition.²⁹⁸ For example, in Germany in 1997, strong competition among insurance companies resulted in falling rates.²⁹⁹

As for more general European insurance rate trends, total premiums for the overall countries represented by the European insurance and reinsurance federation (CEA)³⁰⁰ grew in real terms by 1.2 percent in 2007, compared to an annual increase of 6.5 percent in the two previous years.³⁰¹ The slowdown in the rate of total premium increase was due to

²⁹⁴ Turchetti & Daraio, *supra* note 249, at 205.

²⁹⁵ *Id.* at 207-08.

²⁹⁶ *Id.* at 217.

²⁹⁷ 5 SWISS RE, Sigma, *Upheaval in Insurance Markets – Results Still Good Despite Increased Competition, Forecast for the Biggest non-Life Markets in 1998 and 1999*, at 19 (1998), media.swissre.com/documents/1b477a804659d8e893f8df4ba16c05ab-17_Aug_1998_Upheaval_in_insurance.pdf.

²⁹⁸ *Id.* at 4.

²⁹⁹ *Id.* at 14.

³⁰⁰ The CEA (Comité Européen des Assurances) is the European insurance and reinsurance federation; its members are the national insurance associations in 32 European countries (Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom). The statistical data presented in the text refer to the above-mentioned 32 countries beside Lithuania. COMITÉ EUROPÉEN DES ASSURANCES, <http://www.cea.eu/>.

³⁰¹ 2008 CEA Statistics, *supra* note 280, at 11. This part of the article considers 2007 data since 2008 and 2009 insurance premium data are affected by the impact of the financial crisis. Due to the financial crisis, gross written premiums declined by 6% in 2008. *See CEA Statistics No. 37: European Insurance*

strong competition between insurance companies in the non-life sector.³⁰² In the non-life sector premium growth in 2007 slowed down to 0.4 percent in real terms.³⁰³ In Western Europe, eight out of fifteen markets experienced a decrease in premium volumes.³⁰⁴ For example, in Germany and in the United Kingdom, which are the two largest European non-life markets, premium volume fell respectively 1.4 percent and 0.7 percent respectively.³⁰⁵ The link between the general slowdown in total European non-life premiums and lower insurance rates could also be seen in the motor vehicle insurance line, which is the biggest line of non-life insurance in Europe, accounting for 31 percent of total premiums in 2007.³⁰⁶ Motor insurance premiums declined by 0.4 percent in real terms in 2007 and by 2 percent in 2006.³⁰⁷ This reduction was caused by lower rates due to strong competition between insurance companies.³⁰⁸

Thus, deregulation and the establishment of a single insurance market in Europe had positive effects by intensifying competition among

in *Figures*, COMITÉ EUROPÉEN DES ASSURANCES 9 (Oct., 2009), <http://www.cea.eu/uploads/Modules/Publications/eif-2009.pdf> [hereinafter 2009 CEA Statistics]. In 2009 the European insurance industry weathered the crisis better as to premium growth and total European premiums increased by 2.9%. See *CEA Statistics No. 42: European Insurance in Figures*, COMITÉ EUROPÉEN DES ASSURANCES 10 (Nov., 2010), http://www.cea.eu/uploads/Modules/Publications/1290503264_european-insurance-in-figures.pdf [hereinafter 2010 CEA Statistics].

³⁰² 2008 CEA Statistics, *supra* note 302, at 11.

³⁰³ *Id.* at 14.

³⁰⁴ *Id.*

³⁰⁵ *Id.*

³⁰⁶ See *id.* at 15. See also 2009 CEA Statistics, *supra* note 301, at 14 (showing a decline in European motor insurance premiums in 2008 due to insurers' efforts to improve the value for money of products sold, the strong competition in the market and the decline in new car sales because of the economic crisis); 2010 CEA Statistics, *supra* note 301, at 15 (showing a decline in European motor insurance premiums in 2009 mainly due to the competitiveness of the market and the economic crisis); *Retail Insurance Market Study*, EUROPE ECONOMICS 100, 104-05 (Nov. 26, 2009), http://ec.europa.eu/internal_market/insurance/docs/motor/20100302rim_en.pdf (stating that Europe has the largest motor insurance market in the world, with almost € 119 billion motor insurance premiums in the EU27 in 2008) [hereinafter *Retail Insurance Market Study*].

³⁰⁷ 2008 CEA Statistics, *supra* note 280, at 15.

³⁰⁸ *Id.*

insurers.³⁰⁹ Insurance companies were able to adjust their rates following deregulation, and there were not substantial rate increases.³¹⁰

C. THE CASE FOR RATE DEREGULATION

The concern that deregulation could lead to monopolistic or oligopolistic pricing is controverted by the fact that the U.S. insurance market is competitive and does not require regulation of insurance rates.³¹¹ Table 2 above, for example, shows no evidence of excessive profits by insurers.

Indeed, rate regulation in the U.S. may result in artificially low returns. According to an ISO analysis, the profitability of property and casualty insurers measured under generally accepted accounting principles (GAAP)³¹² is lower than other industries.³¹³ The return on net worth of both large property and casualty insurance companies and the entire property and casualty insurance industry for the period 1983 to 2009 was lower than the return on net worth for the Fortune 500 combined companies except in 1986 and in 1987.³¹⁴ Other industries also had higher rates of return compared to the property and casualty insurance industry over that period.

³⁰⁹ See *Retail Insurance Market Study*, *supra* note 306, at xxi, 89-90 (analyzing the European motor insurance market).

³¹⁰ See *Id.*

³¹¹ See *supra* pp. 134-39.

³¹² The data reported in the annual statement filed by insurance companies with state Insurance Departments and the Internal Revenue Service are on a statutory accounting principles (SAP) basis, that tends to be more conservative than GAAP. Therefore, in order to make comparisons with other industries it is appropriate to consider the adjustment of the insurers' profitability on a GAAP basis. See THE INSURANCE FACT BOOK 2011, *supra* note 229, at 39.

³¹³ THE INSURANCE FACT BOOK 2011, *supra* note 229, at 39.

³¹⁴ See INSURANCE INFORMATION INSTITUTE, THE INSURANCE FACT BOOK 2009 33 (2009); THE INSURANCE FACT BOOK 2011, *supra* note 229, at 39.

Table 3 – 2000-2009 Annual Rate of Return: Net Income After Taxes as a Percent of Equity

Year	Property and Casualty Insurance		Selected Other Industries		Fortune 500 combined industrials and service
	Statutory Accounting	GAAP Accounting	Commercial banks	Electric and gas utilities	
2000	6.2 percent	5.9 percent	16.7 percent	11.8 percent	14.6 percent
2001	-2.0 percent	-1.2 percent	14.0 percent	10.5 percent	10.4 percent
2002	3.0 percent	2.1 percent	17.3 percent	7.9 percent	10.2 percent
2003	8.3 percent	8.8 percent	14.9 percent	10.5 percent	12.6 percent
2004	9.7 percent	9.4 percent	15.5 percent	10.5 percent	13.9 percent
2005	10.9 percent	9.6 percent	16.0 percent	10.0 percent	14.9 percent
2006	14.2 percent	12.7 percent	15.0 percent	11.0 percent	15.4 percent
2007	12.0 percent	10.9 percent	11.0 percent	11.0 percent	15.2 percent
2008	0.8 percent	0.1 percent	3.0 percent	13.0 percent	13.1 percent
2009	6.2 percent	4.7 percent	4.0 percent	9.0 percent	10.5 percent

Source: INSURANCE INFORMATION INSTITUTE, THE INSURANCE FACT BOOK 2011 39 (2011), reprinted with permission. (Source: SNL Financial LC; ISO; Fortune).

Calculating, from data in Table 3, the average rate of return on a GAAP basis for the property casualty insurance industry and the average rate of return for the Fortune 500 combined companies for the period 2000-2009, the return on average for the property and casualty insurers was 6.3 percent and 13.08 percent for the Fortune 500 combined companies.

Table 2 also raises questions about insurers' low profitability.³¹⁵ Although the market growth for commercial lines over the 2000-2009 period indicated that new insurers had incentives to enter the business,³¹⁶ nevertheless many insurers exited the market.³¹⁷ For example, 29 insurers exited the commercial auto liability market, 31 exited the workers'

³¹⁵ See *supra* Table 2.

³¹⁶ It is more appropriate to consider the market growth over the period 2000-2009 than just over the period 2007-2009 since 2007-2009 data may be affected by the impact of the 2008 financial crisis. As to the data for commercial property and casualty insurance before the financial crisis, see *infra* Appendix 2.

³¹⁷ See *supra* Table 2.

compensation market, 32 left the market for inland marine and 43 exited the allied lines market.³¹⁸ The fact that those insurers exited the market may suggest that they did not consider the market profitable enough to remain in business.³¹⁹

It is difficult to establish for certain a direct causal link between this data and insurance rate regulation. At a minimum, rate regulation could be one of the causes depressing insurers' profitability.³²⁰

Rate regulation may affect insurers' profitability since it may limit insurers' ability to adjust their rates according to changes in market conditions. Insurers might need to raise rates when investment income dips or premiums are too low to absorb losses. Yet in prior approval systems insurers may experience delays or denials in getting approval for rate increases. There could also be political pressure on insurance commissioners to keep rates low.³²¹ A commissioner might grant approval for a rate increase lower than that requested by the insurer, either to attain the rate increase over a longer period of time or not at all. This can have adverse effects on insurance companies. Similar concerns surround file and use, use and file, flex rating and modified prior approval systems. With regard to the first two systems, the commissioner retains the right to disapprove the rates filed, while, with regard to the flex rating and modified prior approval system, both require prior approval if the rate change is larger than the specified percentage rate, or if the rate revision is based on a change in expense ratio or rate classifications. Because of the time and expense to meet the rate-filing requirements, insurance companies may have less-than-optimal opportunity to adjust their rates to changes in the

³¹⁸ *See id.*

³¹⁹ In particular, with regard to commercial multiple peril insurance, inland marine and allied lines, more insurers exited the market than entered it (a net loss of 2 in the commercial multiple peril market, of 7 in the inland marine market and of 15 in the market for allied lines). *See supra* Table 2.

³²⁰ *See* Cummins et al., *supra* note 215, at 42-44 (demonstrating by statistical regression analysis that insurance regulation led to significantly lower prices in the majority of states that were regulated during the sample period 1980-1996); Cummins & Harrington, *supra* note 215, at 60 (showing by multiple regression analysis that in competitive rating states loss ratios are significantly lower and average prices significantly higher); Scott E. Harrington, *A Note on the Impact of Auto Insurance Rate Regulation*, 69 REV. ECON. & STAT. 166, 169 (1987) (finding that auto insurance rate regulation increased average loss ratios during the sample period 1976-1981).

³²¹ *See supra* p. 135.

market.³²² While deregulating might result in higher rate volatility, it would permit insurance companies to set appropriate rates in response to changes in market conditions.

Deregulation would also allow insurers to eliminate the costs of complying with rate regulation and prior approval systems. Under the NAIC Property and Casualty Model Law, an insurer has to file with the insurance commissioners “every manual, minimum premium, class rate, rating schedule or rating plan and every other rating rule, and every modification of any of the foregoing which it proposes to use.”³²³ Further, insurers have to submit or incorporate by reference “all supplementary rating and supporting information to be used in support of or in conjunction with a rate,” such as the insurers’ interpretation of statistical data on which they relied, the experience of other insurance companies, and any other relevant information.³²⁴ The commissioner, after reviewing the insurer’s filing, may require that “the insurer’s rates be based upon the insurer’s own loss, special assessment and expense information,” where the insurer’s loss is not actuarially credible, the insurer “may use or supplement its experience with information filed with the commissioner by an advisory organization or statistical agent.”³²⁵ For insurers using the services of an advisory organization, the commissioner may require them to provide “a description of the rationale for such use, including its own information and method of utilization of the advisory organization’s information.”³²⁶

Rate filings, therefore, are a drain on insurers’ time and resources. The process to approve rates can be invasive, lengthy, inaccurate and disputed.³²⁷ Further, a commissioner’s analysis of whether the rates are “excessive, inadequate or unfairly discriminatory” requires a considerable outlay of effort and resources by insurance department staff in order to consider past and prospective loss experience and expenses.³²⁸ The same is

³²² HARRINGTON, *supra* note 1, at 33.

³²³ PROPERTY AND CASUALTY MODEL RATING LAW (PRIOR APPROVAL VERSION) § 5(A)(1) (2009) [hereinafter PRIOR APPROVAL MODEL LAW]; FILE AND USE MODEL LAW, *supra* note 240, at § 6(A)(1).

³²⁴ PRIOR APPROVAL MODEL LAW, *supra* note 323, at § 5(A)(2); FILE AND USE MODEL LAW, *supra* note 240, at § 6(A)(2).

³²⁵ *Prior Approval Model Law*, *supra* note 323, at § 5(A)(4); FILE AND USE MODEL LAW, *supra* note 240, at § 6(A)(4).

³²⁶ *Prior Approval Model Law*, *supra* note 323, at § 5(A)(5); FILE AND USE MODEL LAW, *supra* note 240, at § 6(A)(5).

³²⁷ HARRINGTON, *supra* note 1, at 31.

³²⁸ *Prior Approval Model Law*, *supra* note 323, at § 4(B); FILE AND USE MODEL LAW, *supra* note 240, at § 5(A)(4).

true for a commissioner's determination of whether there is competition in the market.³²⁹ Rate deregulation could allow insurance departments to fully devote themselves to other more important supervisory activities such as solvency supervision.³³⁰

In addition, the effect of rate regulation on the availability of insurance can be seen by analyzing the residual market. Generally, a declining residual market means that insurance is relatively more available in the voluntary market, and vice versa. Thus, it is of concern that residual market shares increase along with the degree of rate regulation. Table 4 compares the size of the voluntary market and the residual market by state for private passenger car insurance for the year 2008.

Table 4 – Private Passenger Cars Insured in the Voluntary and Residual Market, 2008

State	Voluntary Market	Residual Market	Total	Residual market as a percentage of total
Alabama	3,384,021	6	3,384,027	< 0.001 percent
Alaska	437,274	122	437,396	0.028 percent
Arizona	4,130,900	20	4,130,920	< 0.001 percent
Arkansas	2,069,310	0	2,069,310	< 0.001 percent
California	24,127,758	5,941	24,133,699	0.025 percent
Colorado	3,667,061	0	3,667,061	< 0.001 percent
Connecticut	2,442,996	487	2,443,483	0.020 percent
Delaware	608,459	25	608,484	0.004 percent
D.C.	221,678	457	222,135	0.206 percent
Florida	11,288,408	6	11,288,414	< 0.001 percent
Georgia	6,789,526	3	6,789,529	< 0.001 percent
Hawaii	796,742	5,188	801,930	0.647 percent
Idaho	1,068,562	38	1,068,600	0.004 percent
Illinois	7,936,919	1,153	7,938,072	0.015 percent
Indiana	4,578,960	6	4,578,966	< 0.001 percent
Iowa	2,398,138	9	2,398,147	< 0.001 percent

³²⁹ FILE AND USE MODEL LAW, *supra* note 240, at § 4, 8.

³³⁰ See *supra* p. 135.

Kansas	2,349,365	1,327	2,350,692	0.056 percent
Kentucky	3,013,470	64	3,013,534	0.002 percent
Louisiana	2,834,988	7	2,834,995	< 0.001 percent
Maine	1,022,278	28	1,022,306	0.003 percent
Maryland	3,792,401	73,328	3,865,729	1.897 percent
Massachusetts	3,955,971	112,891	4,068,862	2.775 percent
Michigan	6,164,846	1,297	6,166,143	0.021 percent
Minnesota	3,746,861	5	3,746,866	< 0.001 percent
Mississippi	2,076,581	76	2,076,657	0.004 percent
Missouri	4,195,783	41	4,195,824	0.001 percent
Montana	775,934	230	776,164	0.030 percent
Nebraska	1,501,473	4	1,501,477	< 0.001 percent
Nevada	1,793,132	23	1,793,155	0.001 percent
New Hampshire	904,727	710	905,437	0.078 percent
New Jersey	5,290,260	15,048	5,305,308	0.284 percent
New Mexico	1,455,016	24	1,455,040	0.002 percent
New York	9,233,103	92,283	9,325,386	0.990 percent
North Carolina	5,607,617	1,442,470	7,050,087	20.460 percent
North Dakota	592,814	4	592,818	0.001 percent
Ohio	8,029,756	0	8,029,756	< 0.001 percent
Oklahoma	2,719,636	52	2,719,688	0.002 percent
Oregon	2,724,683	9	2,724,692	< 0.001 percent
Pennsylvania	8,483,438	19,151	8,502,589	0.225 percent
Rhode Island	663,890	9,335	673,225	1.387 percent
South Carolina	3,294,512	1	3,294,513	< 0.001 percent
South Dakota	681,839	0	681,839	< 0.001 percent
Tennessee	4,187,461	24	4,187,485	0.001 percent
Texas	Data not available	Data not available	Data not available	Data not available
Utah	1,808,234	2	1,808,236	< 0.001 percent
Vermont	474,881	450	475,331	0.095 percent
Virginia	6,023,910	1,460	6,025,370	0.024 percent
Washington	4,513,296	0	4,513,296	< 0.001 percent

West Virginia	1,305,657	39	1,305,696	0.003 percent
Wisconsin	3,674,130	0	3,674,130	< 0.001 percent
Wyoming	503,741	1	503,742	< 0.001 percent
Nationwide	185,342,396	1,783,845	187,126,241	0.953 percent

Source: INSURANCE INFORMATION INSTITUTE, *THE INSURANCE FACT BOOK 2011* 71-72 (2011), reprinted with permission. (Source: Automobile Insurance Plans Service Office).

Of the states with a higher residual market share relative to the total private passenger cars insured in 2008, North Carolina (20.460 percent), Massachusetts (2.775 percent), and New York (0.990 percent) had strict rate regulation systems for automobile insurance.³³¹ North Carolina and New York had prior-approval rating laws with regard to auto-insurance,³³² while in Massachusetts until April 2008 auto insurance rates were set by the commissioner.³³³ Conversely, some of the states with a lower residual market share were states with less restrictive rating systems like file and use or use and file: Arizona, Arkansas, Colorado, Indiana, Iowa, Minnesota, Nebraska (less than 0.001 percent), Delaware (0.004 percent), Idaho (0.004 percent), Illinois (0.015 percent).³³⁴

This suggests that rate regulation may have negative effects on insurance availability. Rate suppression, especially, may force insurers to tighten underwriting, forcing consumers to turn to the residual markets for coverage.³³⁵ Understandably, insurers will not underwrite higher risk consumers if rates are too low to cover their possible costs. In the worst

³³¹ *But see* the cases of Maryland and Rhode Island having a quite high residual market share (1.897 percent and 1.387 percent respectively) even though they adopt less restrictive rating systems: file-and-use the former and flex-rating the latter.

³³² *See* NAIC, *Auto Insurance Database Report 2005/2006* 231, (2008) [hereinafter *Auto Insurance Database Report*]. In June 2008 the New York legislature approved flex-rating legislation for auto insurance providing that, subject to some conditions, overall average rate level increases or decreases of 5 percent above or below the previously filed rates may take effect without obtaining prior regulatory approval.

³³³ *See supra* pp. 126-27.

³³⁴ With regard to the rating systems for auto insurance adopted in the states, *see Auto Insurance Database Report, supra* note 332, at 231.

³³⁵ Residual market mechanisms are statutory arrangements that permit to provide insurance to people considered ineligible for coverage in the voluntary market.

case rate suppression could cause insurers who cannot offset low rates with decreased costs to exit the market.³³⁶ Consider, for example, the number of insurers who exited California following the introduction of Proposition 103³³⁷ and Massachusetts, New Jersey and South Carolina because of strict rate regulation in the auto insurance market.³³⁸ In particular, New Jersey, before 2003, had had a highly regulated automobile insurance market that prompted over 20 insurers to exit over a period of 10 years.³³⁹ After the state enacted reforms in 2003 increasing competition in the market,³⁴⁰ the number of insurance companies changed from 17 to 39, the availability of insurance increased, and insurance prices fell for most policyholders.³⁴¹ The same occurred in South Carolina where a less restrictive rating law was passed in 1999. Afterwards, the number of insurers offering automobile insurance almost doubled and the residual market share and rate levels fell.³⁴²

Deregulation of rates, therefore, would avoid problems with the availability of insurance by allowing insurers to charge an appropriate price to cover their costs and earn a reasonable profit. Although deregulation might lead to higher rates, the benefits of rate suppression are not worth the cost of restricted availability. While rate suppression may make insurance affordable in the short run, in the long run it will cause insurers to exit from the market with consequent problems for consumers and the social welfare. Instead, rate deregulation will result in appropriate prices for insurance

³³⁶ Harrington, *supra* note 118, at 189.

³³⁷ Editorial, *California Smashup*, WALL ST. J., Nov. 15, 1988, at A22 (discussing the exit of forty insurance companies from California due to the enforcement of Proposition 103 rate rollback).

³³⁸ Harrington, *supra* note 118, at 189.

³³⁹ Tennyson, *supra* note 179, at 16.

³⁴⁰ The Auto Insurance Reform Act approved in June 2003 (P.L. 2003, c. 89), introduced *inter alia* (i) the phase-out and final elimination of the “take-all-comers” provisions of the Fair Automobile Insurance Reform Act of 1990 (P.L. 1990, c. 8); (ii) amendments to the prior approval rate filing provision to establish a time-line for regulatory action; (iii) changes in the expedited rate filing procedure by raising the ceiling for rate increases; (iv) provisions that simplify the procedures to be used by insurers to withdraw from selling a particular type of insurance or to withdraw from the state; (v) measures to combat insurance fraud and provide for consumer protection and education. Further, the 2003 Act also amended the New Jersey’s excess profits law, according to which insurers were prohibited from earning more than 6 percent in profits from the sales of auto insurance policies over a three-year period. The Act extended that period from three to seven years.

³⁴¹ Tennyson, *supra* note 179, at 16.

³⁴² HARRINGTON, *supra* note 1, at 22; Tennyson, *supra* note 179, at 16.

because insurance companies in a competitive market will supply products in the long run at prices equaling their average costs plus a reasonable profit.

Rate deregulation will also lead to more consumer choices because, by increasing the range of prices that insurers can charge, the range of products offered to policyholders should increase as well. Rate regulation limits consumer choice since, to implement a uniform price system, comparable rates must be charged for comparable products. Any effects on the ability of consumers to compare insurance rates can be addressed by increasing disclosure and enhancing regulation of insurance advertising and marketing.³⁴³ In addition, standard policy conditions may also facilitate consumers in comparing insurance policies offered by different insurers.³⁴⁴ This way, policyholders may acquire the knowledge they need to properly compare rates and make informed decisions, taking into account the price, the quality of the policy and the insurer's financial strength. Comparison shopping can be enhanced by the on-line availability of insurance quotations³⁴⁵ and help of independent agents and brokers in assisting policyholders with price comparisons.³⁴⁶

One of the main objectives of rate regulation is to prevent insurance insolvencies that could result from ruinous competition. In the long term a competitive market should reach equilibrium where insurers charge premiums that equal their average costs. Insurance companies like all other enterprises strive to conduct a financially successful business so that they are most unlikely to charge rates not sufficient to cover incurred losses and expenses.

Even if that is not always the case, rate regulation has not avoided insurers' insolvencies. In the United States, around 340 property and

³⁴³ HARRINGTON, *supra* note 1, at 26.

³⁴⁴ *Id.* at 44-45.

³⁴⁵ See, e.g., the Massachusetts Division of Insurance's website on auto insurance premium comparison, http://www.mass.gov/?pageID=ocaterminal&L=4&L0=Home&L1=Consumer&L2=Insurance&L3=Automobile+Insurance&sid=Eoca&b=terminalcontent&f=doi_AutorateCompare_autoratecompare&csid=Eoca. The website gives information on how to contact insurance companies and agents directly for quotes and allows consumers to compare premiums for new private passenger auto insurance across companies for seven policy examples by showing the range of prices and discounts they may qualify for.

³⁴⁶ HARRINGTON, *supra* note 1, at 26-27.

casualty insurance companies became insolvent from 1986 to 2006.³⁴⁷ Furthermore, there have been relatively fewer insolvencies in states with less restrictive rating laws than in those that highly regulate rates.³⁴⁸ Of 79 insolvent insurance companies subject to rate regulation in the period 1946 to 1959, 26 were based in Texas, a state which used to have rate uniformity enforced by law.³⁴⁹ California and Missouri that did not then regulate rates at all had only 3 insolvencies each.³⁵⁰ Similar considerations can be inferred from the data concerning the number of insolvencies in the Italian auto insurance market for the period immediately following deregulation of tariffs.³⁵¹ Rate deregulation, therefore, is consistent with a financially healthy insurance industry. It permits flexibility in the price of insurance and allows insurers to charge appropriate rates in connection with possible market changes and, ultimately, to set rates more aligned with costs, thereby enhancing insurers' financial strength.

CONCLUSION

Rate regulation seems to be based more on an historical tradition than on solid economic arguments. Although deregulation might seem bold in the current financial crisis, it is important to distinguish between the need for rate regulation and the desirability of more effective solvency regulation. Solvency concerns can be addressed by focusing on insurers' reserves and increasing the monitoring of the financial conditions of insurers. For these reasons rate freedom should replace regulation of rates.

³⁴⁷ American Academy of Actuaries, *Risk Focused Surveillance Framework* 14 (2006), http://www.naic.org/documents/committees_c_catf_aaa_risk_focus_ed_surveillance.doc.

³⁴⁸ CRANE, *supra* note 189, at 95.

³⁴⁹ *Id.*

³⁵⁰ *Id.*

³⁵¹ *See supra* p. 148.

Appendix 1

Rate Filing Methods

STATE	FILING METHOD	LINES
Alabama	file and use prior approval	commercial lines, title medical malpractice, personal lines, property and inland marine, casualty and surety, workers' compensation
Alaska	prior approval flex rating on rate changes file and use rate changes	medical malpractice, title workers' compensation, assigned risk rates all property and casualty lines except workers' compensation, medical malpractice, assigned risk all property and casualty lines except workers' compensation, medical malpractice, assigned risk
Arizona	file and use use and file	workers' compensation, title other property and casualty lines
Arkansas	file and use (competitive market); prior approval (non-competitive market) no filing prior approval	personal lines and small commercial risks large commercial risks workers' compensation
California	prior approval file and use	all property and casualty lines Title
Colorado	prior approval file and use no file; must maintain documentation	workers' compensation loss cost filing by a rating organization; auto assigned risk all other property and casualty lines, except exempt commercial policyholders, title exempt commercial policyholders

Connecticut	file and use prior approval	commercial lines (exception), personal lines medical malpractice (for rate increasing 7.5 percent or more over last rates filed), title
Delaware	file and use	all lines except title
District of Columbia	file and use prior approval	all lines workers' compensation and medical malpractice
Florida	file and use or use and file prior approval rate set by the Commissioner	all lines except title and workers' compensation workers' compensation title
Georgia	prior approval file and use no file	personal private passenger auto other property and casualty lines large commercial risks
Hawaii	prior approval	property and casualty lines
Idaho	prior approval use and file	workers' compensation, title other property and casualty lines
Illinois	use and file file and use	private passenger auto, taxicabs, motorcycles, homeowners, allied lines, dwelling fire, liquor liability, workers' compensation medical malpractice, group inland marine
Indiana	file and use modified file and use	property and casualty lines workers' compensation
Iowa	prior approval use and file	workers' compensation, other property and casualty lines, title homeowners, private passenger auto
Kansas	prior approval file and use no file	workers' compensation personal and commercial lines large commercial insured, medical malpractice

Kentucky	use and file (competitive market); prior approval (non-competitive market); prior approval of any rates which when combined with any rating factors effectively change pre-tax premium of any particular policy by more than +/- 25 percent in any 12-month period of time	personal lines, auto guaranty, credit, medical malpractice, workers' compensation
	no file	other commercial lines
	file and use	title
Louisiana	prior approval	all property and casualty lines
	no file	workers' compensation (competitive market)
Maine *	modified file and use	property and casualty lines, title
	no filing	large commercial risks
	prior approval	workers' compensation
Maryland	file and use	lines designated by the commissioner as competitive
	prior approval	property and casualty lines, title
Massachusetts	file and use or set by the commissioner	motor vehicle (filing method based on finding of existence of competitive market by commissioner)
	file and use	all other lines
Michigan	file and use	auto, homeowners, workers' compensation, inland marine, title
	prior approval	property excluding auto and homeowners
Minnesota	file and use	all lines except workers' compensation
	prior approval	workers' compensation
Mississippi	prior approval	property and casualty lines
Missouri	informational filing only	commercial property and casualty lines
	use and file	other property and casualty lines, workers' compensation

Montana	file and use	property and casualty lines, title
Nebraska	file and use prior approval	personal lines, workers' compensation, most commercial lines, crop, professional liability, excess and large deductible workers' compensation medical professional liability, title
Nevada	prior approval prior approval file and use file and use	all personal lines, medical professional liability rates, except surety workers' compensation loss costs and assigned risk rates workers' compensation loss cost multipliers and supplementary rate information title
New Hampshire	file and use (competitive market); prior approval (non-competitive market) use and file prior approval no filing use and file prior approval no filing required	personal lines (competitive market) commercial lines (competitive market), workers' compensation commercial lines (non-competitive market) ocean marine, aircraft, financial guaranty, boiler and machinery title workers' compensation ocean marine, aircraft, financial guaranty, boiler and machinery
New Jersey	use and file prior approval	commercial lines other property and casualty lines, workers' compensation, title
New Mexico	prior approval no file commissioner-set rates	property and casualty lines (non-competitive markets, reverse competitive and residual markets), workers' compensation property and casualty lines (competitive markets except workers' compensation and medical professional liability) title

New York	prior approval	workers' compensation, title, medical malpractice, personal and commercial lines
	flex rating	auto
	file and use	other property and casualty lines
North Carolina	prior approval	personal auto, homeowners
	modified file and use	commercial property and casualty lines
	file and use	workers' compensation, title
North Dakota	prior approval	all lines except workers' compensation and aircraft
Ohio	file and use	all other lines
	file and use (competitive market)	commercial casualty
	prior approval (non-competitive market)	medical malpractice
	file and use	
Oklahoma	use and file	property and casualty lines (competitive market)
	file and use	property and casualty lines (non-competitive market), medical malpractice
Oregon	flex rating	commercial casualty
	prior approval	workers' compensation, title
	file and use	other property and casualty lines
Pennsylvania	prior approval	property and casualty lines
	exempt from filing	large commercial risks
	file and use	small commercial risks
Rhode Island	file and use	casualty, property, title
	prior approval	workers' compensation
	no file	large commercial risks
	flex rating	casualty insurance, fire and marine
South Carolina	prior approval	all lines
	prior approval or file and use	commercial auto rate changes of 7 percent or less

South Dakota	no file	large commercial risks
	file and use	all lines except title
	prior approval	title
Tennessee	prior approval	personal lines, workers compensation
	use and file	commercial lines, workers compensation loss cost multipliers
	file and use	title
Texas	file and use	all lines
	commissioner sets rates	title
Utah	use and file	property and casualty lines
	file and use	title, workers' compensation
Vermont	prior approval	workers' compensation, auto (assigned risk), property and casualty lines (non-competitive market)
	use and file	property and casualty lines (except claims made and assigned risk), title and other types of workers' compensation (voluntary market)
Virginia	prior approval	residual market for workers' compensation and automobile; home protection, credit property, credit involuntary unemployment
	no file	large commercial risks, title
	file and use (competitive market)	general liability, homeowners, fire, miscellaneous property and casualty, boiler and machinery, surety, credit, inland marine, farm owners', mortgage guaranty commercial multi-peril; professional liability and legal services
	60 days prior filing requirement for non-competitive lines	property and casualty lines identified by commissioner after hearing

Washington	prior approval	property and casualty (except commercial lines), medical malpractice, workers' compensation commercial lines
	use and file	
	file and use	title
West Virginia	prior approval	other property and casualty lines, excluding workers' compensation commercial lines
	file and use	
Wisconsin	use and file	property and casualty, title
	prior approval	workers' compensation
Wyoming	prior approval	title, medical malpractice
	no filing (competitive market); prior approval (non-competitive market)	property and casualty

SOURCE: NAIC, 2 COMPENDIUM OF STATE LAWS ON INSURANCE TOPICS, HEALTH/LIFE/PROPERTY/CASUALTY II-PA-10-1–II-PA-10-20 (2010), reprinted with permission. The NAIC does not endorse any analysis or conclusions based on use of its data; NAIC, 2 COMPENDIUM OF STATE LAWS ON INSURANCE TOPICS, HEALTH/LIFE/PROPERTY/CASUALTY II-PA-10-9 (2008), reprinted with permission. The NAIC does not endorse any analysis or conclusions based on use of its data.

Appendix 2

2007 Commercial Lines Data – Countrywide

LINE OF BUSINESS	PREMIUMS WRITTEN	MARKET SHARES FOUR LARGEST GROUPS	HHI BASED ON PREMIUM	NUMBER OF SELLERS (GROUPS)	NUMBER OF ENTRIES LAST 5 YEARS	NUMBER OF EXITS LAST 5 YEARS	MARKET GROWTH LAST 3 YEARS	MARKET GROWTH LAST 10 YEARS	RETURN ON NET WORTH 10-YEAR MEAN
Commercial Auto Liability	22,071,577,526	27.64 percent	308.6	103	28	35	3.50 percent	62.76 percent	6.45 percent
Commercial Auto Physical	7,255,767,155	24.45 percent	260	114	24	29	6.44 percent	34.54 percent	11.72 percent
Commercial Auto Total	29,327,344,681	25.89 percent	281.5	109	27	33	4.24 percent	54.73 percent	7.39 percent
Commercial Multiple Peril	36,138,799,711	26.75 percent	327.8	105	36	34	4.52 percent	68.36 percent	7.00 percent
Fire	12,235,775,465	38.74 percent	579.2	95	32	29	26.33 percent	152.40 percent	17.39 percent
Allied Lines	11,399,110,261	42.38 percent	621.6	85	29	37	48.72 percent	242.60 percent	8.41 percent
Inland Marine	15,408,763,035	33.44 percent	450	78	25	30	21.02 percent	111.90 percent	18.46 percent
Medical Malpractice	11,684,425,721	25.27 percent	289.9	99	41	25	4.24 percent	87.39 percent	6.15 percent
Other Liability	54,845,528,333	38.39 percent	627.9	88	36	25	1.28 percent	126.30 percent	5.28 percent
Workers Compensation	52,247,498,240	34.20 percent	425.5	104	35	37	5.77 percent	71.94 percent	6.76 percent

SOURCE: NAIC, 2007 Commercial Lines Competition Database Report 13 (2008), reprinted with permission. The NAIC does not endorse any analysis or conclusions based on use of its data.

ERISA: REMEDIES, PREEMPTION AND THE NEED FOR MORE STATE REGULATORY OVERSIGHT

JILLIAN REDDING, ESQ.*

This article explores the current state of ERISA law and its effects concerning good faith, fiduciary breaches, and the remedies available under ERISA. Recent case law provides that the duty of good faith applies in an ERISA context; however, any breaches result in recovery to the employee benefit plan and not usually to the injured victim. According to case law, ERISA precludes state remedies, even laws specific to insurance. In part one, this paper provides an overview of ERISA and why state remedies or more state oversight are necessary to protect beneficiaries. Section two discusses the legislative background of ERISA. Section three discusses several cases that illustrate ERISA's lack of appropriate remedies for fiduciary breaches. Section four provides a case study of Unum Provident, a disability insurer that made such egregious breaches that it was required to have long-term, strict oversight by the state departments of insurance. The last section compares several theories by other authors on how to improve ERISA's remedies, along with this author's argument that the most appropriate and efficient way to remedy breaches is to require strict oversight by the state departments of insurance, which has proved to be most successful in the Unum Provident case.

I. INTRODUCTION

In March 2010, Congress passed the Patient Protection and Affordable Care Act, a health care bill that expands health care to a greater portion of Americans and prohibits insurers from rejecting applicants based on, among other things, pre-existing conditions.¹ This law was passed in

* LLM., Insurance Law, University of Connecticut School of Law, May 2010; J.D., University of Connecticut School of Law, January 2010; B.A., University of Massachusetts, June 2006. Sincere thanks to Professor Kochenburger for his mentoring for this article. I owe a debt of gratitude to Jared Cantor for repeatedly reading this article and for listening to me talk about insurance law when no one

part because of the severe ongoing crises in healthcare. In 2009 alone, the U.S. is estimated to have spent 17.3% of the gross domestic product on healthcare.² This notable day will go down in history along with days that other great social welfare bills were passed, such as Social Security and Medicare. One glaring oversight, however, is that the bill fails to correct an erroneous interpretation by the courts that has been ongoing for decades: the remedies available under the Employee Retirement Income Security Act (ERISA).³

This paper concerns ERISA, the statutory remedies available to participants and beneficiaries, and a way to protect insureds from egregious insurer behavior. ERISA law governs the benefit plans offered by employers, such as pension and healthcare plans (called “welfare benefit plans”). The Supreme Court has held that the remedies specifically enumerated in section 502 of ERISA are the exclusive remedies available to participants and beneficiaries in any claims relating to ERISA, and has interpreted the “other appropriate equitable relief” provision to preclude make-whole relief, such as money damages, for serious fiduciary breaches. The Supreme Court interprets ERISA’s remedies section narrowly instead of following the true intent of ERISA: protecting the benefits of participants and beneficiaries of such plans balanced against the need to encourage employers to offer such plans. Most courts have decided that, in the competing interests of protecting the participants versus less regulation for employers, Congress intended for less regulation to encourage employers to offer plans. This results in less protection for breaches of fiduciary duties. In doing so, the courts reject that any state tort claims may

else found it interesting, and for being a fantastic fiancée. I am also grateful to my family and friends for their unending support, especially my parents, Lorraine and Lawrence Redding, and my grandparents, Bill and Sandy Watson.

¹ See Patient Protection and Affordable Care Act, Pub. L. No. 111-148, 124 Stat. 119 (2010).

² NHE Fact Sheet, Centers for Medicare and Medicaid Services, NATIONAL HEALTHCARE EXPENDITURE PROJECTIONS FOR 2009-2010, http://www.cms.gov/NationalHealthExpendData/25_NHE_Fact_Sheet.asp (last visited May 16, 2010); *see also* Trends in Health Care Costs and Spending, The Kaiser Family Foundation Report on Trends in Health Care Costs and Spending (March 2009), available at http://www.kff.org/insurance/upload/7692_02.pdf (estimating that the U.S. spent 17.6% of GDP on healthcare) (last visited May 10, 2010).

³ See Employee Retirement Income Security Act of 1974, Pub. L. No. 93-406, 88 Stat. 829 (codified as amended at 29 U.S.C. § 1001 (1988)).

apply to breaches of fiduciary duties in employee benefit health plans, such as insurance unfair trade practices, bad faith, or negligence.⁴

Section two of this paper discusses the history and Congressional intent in passing ERISA. Section three illustrates the problems of ERISA's lack of remedies for beneficiaries by summarizing major case law, legal articles, and investigative journalism by the popular show *Good Morning America*. Section four is a case study of a major disability insurer, Unum Group,⁵ and the litigation and ultimate regulatory oversight due to egregious bad faith on the part of the company. Section five argues for more regulatory oversight for ERISA insurers, which has proved successful in the Unum example and arguably changed the company into a possible model of the industry.

Ultimately, this paper supports the viewpoints of several writers, namely that trust law supports consequential relief in certain instances and state insurance unfair practices statutes should be saved from preemption under the savings clause, including its remedies, as Congress expressly stated that such insurance laws must apply to employee benefit plans.⁶ However, as Congress and the courts are unlikely to change, it is up to the regulators – the state departments of insurance – to effect this protection of

⁴ It is of note that penalties for violations of such claims are what keep insurance companies in check: a fear of large monetary penalties by the courts. This check is completely lacking in the ERISA context because of ERISA's conflict preemption in section 514, which requires ERISA's regulations to "supersede any and all State laws insofar as they may now or hereafter relate to any employee benefit plan." See Employee Retirement Income Security Act of 1974, 29 U.S.C. § 1144(a) (2007).

⁵ Unum Group is the parent company of Unum Life Insurance Company, the Paul Revere Life Insurance Company, Provident Life and Accident Insurance Company, Provident Life and Casualty Insurance Company, and First Unum Life Insurance Company.

⁶ See Donald Bogan, *ERISA: The Savings Clause, § 502 Implied Preemption, Complete Preemption, and State Law Remedies*, 42 SANTA CLARA L. REV. 105, 115 (2001); John H. Langbein, *Trust Law As Regulatory Law: The Unum/Provident Scandal and Judicial Review of Benefit Denials Under ERISA*, 101 NW. U. L. REV. 1315, 1340-41 (2007); John H. Langbein, *What ERISA Means by "Equitable": The Supreme Court's Trail of Error in Russell, Mertens, and Great-West*, 103 COLUM. L. REV. 1317, 1338, 1362 (2003); Charlotte Johnson, *Justice Ginsburg's Fiduciary Loophole: A Viable Achilles' Heel to HMOs' Impenetrable ERISA Shield*, 2006 BYU L. REV. 1589, 1591 (2006); Susan Harthill, *A Square Peg in a Round Hole: Whether Traditional Trust Law "Make-Whole" Relief is Available Under ERISA Section 502(a)(3)*, 61 OKLA. L. REV. 721, 722-24 (2008).

its citizens from bad faith practices by ERISA insurers. The purpose of this paper is to encourage the state departments of insurance to promulgate, enact, and enforce regulations for practices and procedures for ERISA insurers to follow, as they did for Unum in the Regulatory Settlement Agreements. The regulations should focus on the duties of good faith and fair dealing that insurers owe insureds, but is currently lacking in the ERISA landscape.

II. THE HISTORY OF ERISA

ERISA was passed in 1974 with the intent to protect “millions of employees and their dependents” in their retirement benefits, as stated in the Congressional findings and declaration of policy.⁷ This law established a regulatory and guaranty system designed to ensure that employees received the retirement benefits promised. However, some note that the health insurance provisions were hastily added last minute.⁸

ERISA was a reaction to the previously enacted Welfare and Pension Plans Disclosure Act (WPPDA), which sought to provide some federal oversight of retirement benefits.⁹ WPPDA failed to provide the necessary mechanisms to adequately protect employee retirement benefits, as it lacked a method to control administration of the plans or a way to remedy abuses in plan administration, and retired workers lost anticipated benefits promised by employers.¹⁰ Recognizing this failure, the Senate Committee on Labor and Public Welfare appointed a subcommittee to investigate the problem, concluded that WPPDA lacked the necessary substantive regulatory controls, and suggested a that new “comprehensive and reticulated statute” be enacted to correctly regulate the pension industry.¹¹ This new regulation was to protect against, as one writer termed

⁷ 29 U.S.C. § 1001; ERISA § 2.

⁸ TOM BAKER, INSURANCE LAW AND POLICY 130 (2d ed. Aspen Publishers 2008).

⁹ Welfare and Pension Plans Disclosure Act, Pub. L. No. 85-836, 72 Stat. 997, *repealed by* ERISA § 111(a)(1), 29 U.S.C. § 1031(a)(1) (1974).

¹⁰ Bogan, *supra* note 6. Notably, the WPPDA preserved state remedies to supplement its regulation. *See* Act of Mar. 20, 1962, Pub L. No. 87-420, 76 Stat. 35 (*repealed* 1974); Bogan, *supra* note 6.

¹¹ Bogan, *supra* note 6, at 115-16 (citing S. REP. NO. 93-127, at 4, *reprinted in* 1974 U.S.C.C.A.N. 4841, *and* 1 SUBCOMM. ON LABOR OF THE SENATE COMM. ON LABOR AND PUB. WELFARE, 94TH CONG., LEGISLATIVE HISTORY OF THE EMPLOYEE RET. INCOME SEC. ACT OF 1974 at 590 (Comm. Print 1976) [hereinafter

it, default risk and administration risk.¹² Default risk pertains to the danger that an employer may dishonor the promised pension, and applies mainly to defined benefit pension plans.¹³ Administration risk is the danger that the fiduciary (person responsible for managing and investing plan assets and/or paying claims) may abuse his or her authority by performing inappropriately, misusing plan assets, or improperly refusing to pay promised benefits.¹⁴ The health care and disability insurance issue concerns this latter risk because it is where the least protection is provided to claimants.

A. CONGRESSIONAL INTENT

In passing ERISA, Congress noted that there was a lack of transparency to employees and adequate safeguards concerning plan operation, thus “it is desirable in the interests of employees and their beneficiaries, and to provide for the general welfare and the free flow of commerce, that disclosure be made and safeguards be provided with respect to the establishment, operation, and administration of such plans.”¹⁵ Congress was very specific in its intentions when passing ERISA and intended “to alleviate certain problems which tend to discourage the maintenance and growth of multiemployer pension plans”; further “provide reasonable protection for the interests of participants and beneficiaries of financially distressed multiemployer pension plans”¹⁶ and “encourage the maintenance and growth of single-employer defined benefit pension plans.”¹⁷ This encouragement illustrates Congress’ intent to provide a uniform framework of regulation for employee benefit plans to lessen the burden of compliance on employers and entice them to offer such plans.¹⁸ Congress also sought to protect the plan participants’ interests by “providing . . . appropriate remedies, sanctions, and ready access to Federal courts.”¹⁹

“Legislative History”]; H.R. REP. NO. 93-533, at 4 (1978), *reprinted in* 1978 U.S.C.C.A.N. 4642, *and in* 2 Legislative History at 2351)).

¹² Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1323.

¹³ *Id.*

¹⁴ *Id.*

¹⁵ 29 U.S.C. § 1001(a) (1988).

¹⁶ *Id.* §§ 1001a(c)(2)-(3).

¹⁷ *Id.* § 1001b(c)(2).

¹⁸ *See Aetna Health, Inc. v. Davila*, 542 U.S. 200, 208 (2004).

¹⁹ 29 U.S.C. § 1001(b).

B. STATUTORY LANGUAGE AND PREEMPTION

ERISA is frustrating to insureds and claimants attempting to utilize state insurance bad faith laws because of its strict preemption laws, which provide an easy out for plan fiduciaries: removal to federal court, thereby preempting all state laws. Under the Supremacy Clause, federal laws may preempt or take precedence over state laws by express provision, implication, or when there is a conflict between federal and state law.²⁰ ERISA has three main provisions that control the preemption of state laws.

The first, the “preemption clause”, in section 514(a), states: “[e]xcept as provided in subsection (b) [the “savings clause”] of this section, the provisions of this subchapter and subchapter III of this chapter shall supersede any and all State laws insofar as they may now or hereafter relate to any employee benefit plan”²¹ The next provision, “the savings clause”, in section 514(b)(2)(A) provides: “[e]xcept as provided in subparagraph (B) [the deemer clause], nothing in this subchapter shall be construed to exempt or relieve any person from any law of any State which regulates insurance, banking, or securities.”²² The last clause, called the “deemer clause”, in section 514(b)(2)(B), has been held only to apply to self-insured plans, and states:

Neither an employee benefit plan . . . nor any trust established under such a plan, shall be deemed to be an insurance company or other insurer, bank, trust company, or investment company or to be engaged in the business of insurance or banking for purposes of any law of any State purporting to regulate insurance companies, insurance contracts, banks, trust companies, or investment companies.²³

²⁰ See U.S. CONST. art IV; N.Y. State Conference of Blue Cross & Blue Shield Plans v. Travelers Ins. Co., 514 U.S. 645, 654 (1995) (stating the supremacy of federal laws over state laws due to the Supremacy Clause, and its effect on ERISA).

²¹ Pilot Life Ins. Co. v. Dedeaux, 481 U.S. 41, 44-45 (1987) (citing 29 U.S.C. § 1144(a)).

²² *Id.* (citing 29 U.S.C. § 1144(b)(2)(A)).

²³ *Id.* (citing 29 U.S.C. § 1144(b)(2)(B)). In *FMC Corp. v. Holliday*, 498 U.S. 52, 61 (1990), the Court held that the deemer clause exempts self-funded ERISA plans from state laws that “regulate insurance” within the meaning of the savings clause.

To put it more clearly, the preemption clause holds that all state insurance laws that apply an ERISA plan are preempted because they “relate” to an ERISA plan. “Relates to” has been an issue much discussed in case law, and in 1995 the U.S. Supreme Court applied a narrower reading than before in determining what exactly “relates to” means.²⁴ In essence, the Court held that if the state law references or targets an ERISA plan, or has a connection to or directly affects the ERISA plan, it is preempted.²⁵ However, the savings clause will “save” state laws that specifically pertain to insurance, banking, or securities, and those laws will still apply to ERISA plans. An example of this is when an employer purchases insurance for group coverage of its employees under a plan, thus uses an insurer, and the plan is subject to state regulation because the employer is using direct insurance and there is an insurance contract.

Under the deemer clause, self-funded plans are *not* subject to state insurance laws. A plan is self-funded when the employer completely funds the plan, or creates a trust for the employee health plans and deposits money for the claims into the trust. It is not considered insurance because there is no actual insurance contract that is transacted by the employer with regards to the plan, and the state law cannot regulate it via its power to regulate insurance contracts.²⁶ Common sense dictates that, due to numerous state insurance regulation laws which are expensive to comply with, employers are more likely to avoid the costs of complying with the insurance state regulations by creating a trust.

When presenting the bills for ERISA, the preemption provisions were described as a “reservation to Federal authority [of] the sole power to regulate the field of employee benefit plans as ERISA’s crowning achievement” by Representative Dent.²⁷ Another politician, Senator Harrison Williams, provided that this provision “and its narrow exceptions, are intended to preempt the field for Federal regulations, thus eliminating

²⁴ See *N.Y. State Conference of Blue Cross & Blue Shield Plans*, 514 U.S. at 656-57.

²⁵ *Id.*

²⁶ See *Holliday*, 498 U.S. at 64-65 (stating that “[o]ur interpretation of the deemer clause makes clear that if a plan is insured, a State may regulate it indirectly through regulation of its insurer and its insurer’s insurance contracts; if the plan is uninsured, the State may not regulate it”). The Court also noted that it realized it was making a distinction between insured and uninsured plans, thus “leaving the former open to indirect regulation while the latter is not.” *Id.* at 62 (quoting *Metro. Life Ins. Co. v. Massachusetts*, 471 U.S. 724, 747 (1985)).

²⁷ *Pilot Life Ins. Co. v. Dedeaux*, 482 U.S. 41, 46 (1987) (citing 120 Cong. Rec. 29197 (1974)).

the threat of conflicting or inconsistent State and local regulation of employee benefit plans. "This principle is intended to apply in its broadest sense to all actions of State or local governments, or any instrumentality thereof, which have the force or effect of law."²⁸ The result: this legislation effectively deregulated employee health and disability benefits by allowing employee benefit plans to largely be exempt from state insurance regulation. While the insurance forms must still be approved by the state department of insurance and the insurer must abide by the funding requirements, the insurer is completely exempt from state unfair insurance practices statutes.

C. REMEDIES UNDER ERISA

Section 502 provides the civil enforcement language for suits brought against plan fiduciaries. It allows participants and beneficiaries to bring suit to: (1) recover benefits due under the plan; (2) enforce his or her rights under the terms of the plan; (3) clarify his or her rights to future benefits under the terms of the plan; or (4) to enjoin fiduciaries from acts or practices that violate ERISA.²⁹ It also allows for "appropriate relief", which includes language from section 1109, which gives the court discretion to order "such other equitable or remedial relief as the court may deem appropriate" for fiduciary breaches.³⁰ It also allows the Secretary of Health to assess and collect civil penalties for certain violations.³¹ This section applies to both employee insurance benefit plans and trusts. Section 502, which carves out jurisdiction for federal district courts in ERISA-related claims, also allows for concurrent jurisdiction (i.e., plaintiff can bring the suit in either state or federal court) to recover plan benefits, enforce benefit rights, or clarify future benefits.³² However, a cause of action will always be subject to removal to federal court.

There is a lack of state remedies and compensatory remedies afforded to plan participants and beneficiaries when an administrator violates the fiduciary duties. As noted above, section 502(a) establishes the remedies a participant or beneficiary may seek when a violation of fiduciary obligation under ERISA occurs, and the remedies are largely remedial in nature. Notably, courts have held that no remedies are available

²⁸ *Id.* (citing 120 Cong. Rec. 29933 (1974)).

²⁹ 29 U.S.C. § 1132(a)(1)(B) (2000).

³⁰ *Id.* (referencing 29 U.S.C. § 1109(a)).

³¹ *Id.*

³² *See* 29 U.S.C. §§ 1132(a)(1)(B), (e).

to plan participants or beneficiaries *unless* it is enumerated in this section, because “Congress clearly expressed an intent that the civil enforcement provisions of ERISA § 502(a) be the exclusive vehicle for actions by ERISA-plan participants and beneficiaries asserting improper processing of a claim for benefits, and that varying state causes of action for claims within the scope of § 502(a) would pose an obstacle to the purposes and objectives of Congress.”³³ There is no ability to bring a bad faith claim against a plan or its fiduciary or be awarded punitive damages.³⁴ This leaves plan participants and beneficiaries in the untenable condition of fighting with the plan’s administrators over benefits, with little to no recourse from the courts, especially if the plan language affords the administrator with discretionary authority to interpret the plan, raising the level of review by a court to arbitrary and capricious.³⁵

In short, if a plaintiff brings suit because of a mishandled benefit claim and suffers health damages due to the denied treatment, all he may receive from the court is (1) an injunction to stop the insurer from wrongfully denying benefits in the future and (2) “other equitable relief.”³⁶ The courts have interpreted the “equitable remedy” to mean that the plaintiff must receive the benefit that was wrongfully denied. However, this does not allow the plaintiff to receive any damages for his suffering, the delay, or for the consequential further treatment required from the wrongful benefit refusal. As one writer states, “The courts have . . . interpreted ERISA’s ‘equitable relief’ provision to prevent an insured from obtaining ‘make-whole relief.’ Make-whole relief includes expenses that an insured may have incurred due to the wrongful denial of benefits, such as physical harm or suffering.”³⁷ The Court’s interpretation of “other equitable relief” defies insurance principles. Under traditional insurance law, plaintiffs are usually entitled to damages for insurer breaches under the notion that the insurance contract is one of adhesion, the insured is not sophisticated and is unable to negotiate in any way with the insurer, and to oppose insurer

³³ See *Aetna Health Ins. v. Davila*, 542 U.S. 200, 214 (2004); *Pilot Life Ins. Co.*, 481 U.S. at 52.

³⁴ *Davila*, 542 U.S. at 214; *Pilot Life Ins. Co.*, 481 U.S. at 52-53; *Mass. Mut. Life Ins. Co. v. Russell*, 473 U.S. 134, 147 (1985).

³⁵ Discretionary authority in plan documents was highlighted and basically approved by the Supreme Court in *Firestone Tire and Rubber Co. v. Bruch*, 489 U.S. 101 (1989).

³⁶ 29 U.S.C. § 1132(a).

³⁷ Elizabeth Khoury, *HMO Liability After Aetna Health, Inc. v. Davila: Are Patients’ Rights At Risk?*, 91 IOWA L. REV. 1621, 1631 (2006) (footnotes and citations omitted).

opportunism.³⁸ Courts utilize contract law, essentially the principles of reasonable expectations and unjust enrichment, to remedy plaintiffs.

D. ERISA, TRUST LAW, AND FIDUCIARY DUTIES

An administrator of an employee benefit plan or trust is considered a fiduciary and manages the assets within the plan or trust for the claims of the participants or beneficiaries. ERISA states the fiduciary duties owed by those who have control over the assets, management and administration of a plan in ERISA sections 404 and 405. First, fiduciaries have the duty of loyalty to the participants and beneficiaries, and must always discharge his or her duties *solely* in the interest of those participants and beneficiaries.³⁹ Next, the fiduciary must act “with the care, skill, prudence, and diligence under the circumstances then prevailing that a prudent man acting in a like capacity and familiar with such matters would use”⁴⁰ Further, the fiduciary must reasonably divest the investment of assets and follow the terms of the ERISA plan, unless it violates ERISA.⁴¹

The duties of loyalty and prudence have the most teeth, as they require the fiduciary to solely act for the participants in a prudent manner – thus the fiduciary must “deal fairly and honestly with beneficiaries”.⁴² Congress based ERISA plan administrators’ fiduciary obligations on trust law.⁴³ This “fiduciary law of plan administration governs claims administration as well as the administration of plan assets.”⁴⁴

³⁸ See *C&J Fertilizer, Inc. v. Allied Mut. Ins. Co.*, 227 N.W. 2d 169 (Iowa 1975). Opportunism is described as the insurer’s ability to refuse to pay claims after the insured has faithfully paid all premiums, and it is too late for the insured to switch insurers. See *supra* note 10.

³⁹ 29 U.S.C. § 1104(a)(1)(A) (2006).

⁴⁰ *Id.* § 1104(a)(1)(B).

⁴¹ *Id.* §§ 1104(a)(1)(C)-(D).

⁴² *Varity v. Howe*, 516 U.S. 489, 506 (1996) (citing GEORGE GLEASON BOGERT & GEORGE TAYLOR BOGERT, *LAW OF TRUSTS AND TRUSTEES* § 543, p. 218-19 (rev. 2d ed. 1992); 2A AUSTIN WAKEMAN SCOTT & WILLIAM FRANKLIN FRATCHER, *THE LAW OF TRUSTS* § 170, at 311-12 (4th ed. 1987); RESTATEMENT (SECOND) OF TRUSTS § 170 (1959)).

⁴³ *Varity*, 516 U.S. at 506; see also Langbein, *Trust Law as Regulatory Law*, *supra* note 6, at 1316 (citing *Firestone Tire and Rubber Co. v. Bruch*, 489 U.S. 101, 115 (1989)); John A. Pereira, *A Fiduciary’s Right to Contribution or Indemnity Under ERISA*, 21 OKLA. CITY U. L. REV. 507, 518 (1996) (citing S. REP. NO. 93-127, at 29 (1973), reprinted in 1974 U.S.C.C.A.N. 4838, 4869).

⁴⁴ Langbein, *Trust Law as Regulatory Law*, *supra* note 6, at 1326.

ERISA subjects the administrators of the plan to a modified and constricted version of the “core substantive rules of trust fiduciary law”.⁴⁵ The plan administrator or fiduciary is expected to primarily focus on ensuring that the benefit recipient's expectations are fulfilled; they should not be primarily focused on protecting the plan's assets.⁴⁶ ERISA allows a fiduciary to be personally liable for a breach of duties, responsibilities, or obligations.⁴⁷ ERISA provides that the fiduciary must “make good to such plan any losses to the plan resulting from each such breach, and to restore to such plan any profits of such fiduciary” made through the use of plan assets.⁴⁸ Importantly, the statute gives the court discretion to order “such other equitable or remedial relief as the court may deem appropriate” for fiduciary breaches.⁴⁹

Under traditional trust law, the trustee is the guardian of the trust's assets.⁵⁰ Trust law recognizes the need to preserve the assets in the trust to satisfy future and present claims, and requires the trustee to “take an impartial account of the interests of all beneficiaries.”⁵¹ When a trustee is given discretion as to his exercise of power, the court may only intercede when there is an abuse of that discretion.⁵² Further, one treatise suggests that a court may remove a trustee's power of discretion when there is a reason to believe that he will not act fairly, such as by showing that the trustee has already acted in bad faith.⁵³ Another source indicates that the

⁴⁵ *Id.* at 1326 (citing 26 U.S.C. § 404(a)(1)(A) (2006), § 1104(a)(1)(A)(i)).

⁴⁶ Pereira, *supra* note 43, at 518 (citing *Pompano v. Michael Schiavone & Sons*, 680 F.2d 911, 914 (2d Cir. 1982) (noting that fiduciary's duty is to insure the honest administration of financially sound plans); H.R. REP. NO. 93-533, at 21-22 (1973), *reprinted in* 1974 U.S.C.C.A.N. 4639, 4659-60 (noting that it is not the fiduciary's duty to primarily focus on the protection of assets)).

⁴⁷ 29 U.S.C. § 1109(a) (2006).

⁴⁸ *Id.*

⁴⁹ *Id.*

⁵⁰ RESTATEMENT (SECOND) OF TRUSTS § 176 (1957).

⁵¹ *Varity Corp. v. Howe*, 516 U.S. 489, 514 (1996) (citing RESTATEMENT (SECOND) OF TRUSTS §§ 183, 232 (1957)).

⁵² RESTATEMENT (SECOND) OF TRUSTS § 187 (1957).

⁵³ *Conkright v. Frommert*, 130 S. Ct. 1640, 1647 (2010) (“If the trustee's failure to pay a reasonable amount [to the beneficiary of the trust] is due to a failure to exercise [the trustee's] discretion honestly and fairly, the court may well fix the amount [to be paid] itself. On the other hand, if the trustee's failure to provide reasonably for the beneficiary is due to a mistake as to the trustee's duties or powers, and there is no reason to believe the trustee will not fairly exercise the discretion once the court has determined the extent of the trustee's duties and powers, the court ordinarily will not fix the amount but will instead direct the

court may intercede when the trustee acts “beyond the bounds of reasonable judgment.”⁵⁴ Yet another treatise states that, after a trustee has abused his discretion, the court may decide “for the trustee how he should act, either by stating the exact result it desires to achieve, or by fixing some limits on the trustee's action and giving him leeway within those limits.”⁵⁵ Further, when there is a conflict, trust law allows a court to scrutinize conflicted discretionary acts.⁵⁶

Notably, trust law allows make-whole relief, including consequential relief, for acts of “negligence or misconduct in the making or retaining of investments.”⁵⁷ The make-whole standard under trust law restores the victim to the position he would have been in had no breach occurred, and includes an award of monetary damages.⁵⁸

Appearing to neglect the trust law precedent, the U.S. Supreme Court has interpreted “other appropriate equitable relief” to be similar to Title VII of the Civil Rights Act of 1964, and declared that ERISA

trustee to make reasonable provision for the beneficiary's support.”) (citing 3 A. SCOTT, W. FRATCHER, & M. ASCHER, SCOTT AND ASCHER ON TRUSTS § 18.2.1, at 1348-49 (5th ed. 2007)) (footnotes omitted).

⁵⁴ *Id.* at 1648 (citing RESTATEMENT (SECOND) OF TRUSTS § 187, cmt. i, at 406 (1957)).

⁵⁵ *Id.* at 1648 (citing GEORGE GLEASON BOGERT & GEORGE TAYLOR BOGERT, LAW OF TRUSTS AND TRUSTEES § 560, at 223 (2d rev. ed. 1980)).

⁵⁶ *Metro. Life Ins. Co. v. Glenn*, 128 S. Ct. 2343, 2349 (2008) (citing RESTATEMENT (SECOND) OF TRUSTS § 107, cmt. f (1957) (discretionary acts of trustee with settlor-approved conflict subject to “careful scrutiny”); RESTATEMENT (SECOND) OF TRUSTS § 107, illus. 1 (1957) (conflict is “a factor to be considered by the court in determining later whether” there has been an “abuse of discretion”); RESTATEMENT (SECOND) OF TRUSTS, § 187, cmt. d (1957); 3 A. SCOTT, W. FRATCHER & M. ASCHER, SCOTT AND ASCHER ON TRUSTS § 18.2, at 1342-43 (5th ed. 2007) (hereinafter SCOTT) (same). *See also, e.g.*, BOGERT § 543, at 264 (rev. 2d ed. 1992) (settlor approval simply permits conflicted individual to act as a trustee); BOGERT § 543(U), at 422-31 (same); SCOTT § 17.2.11, at 1136-39 (same)).

⁵⁷ Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1337 (quoting BOGERT, THE LAW OF TRUSTS AND TRUSTEES § 862 at 38-39 (rev. 2d ed. 1995)).

⁵⁸ *Id.* at 1335 (quoting 3 AUSTIN WAKEMAN SCOTT & WILLIAM FRANKLIN FRATCHER, THE LAW OF TRUSTS § 205, at 237 (4th ed. 1988)); *Great-West Life & Annuity Ins. Co. v. Knudson*, 534 U.S. 204, 214 (2002) (citing RESTATEMENT OF RESTITUTION § 160, cmt. a (1936); GEORGE E. PALMER, LAW OF RESTITUTION § 1.4, at 17, § 3.7, at 262 (1978)).

precludes “awards for compensatory or punitive damages.”⁵⁹ The Court reasoned that such relief must be limited to the scope of appropriate relief typically available in equity.⁶⁰ However, in an amicus brief submitted by the Solicitor General, the argument was made that “equitable relief” included damages claims, because traditionally money damages were available in equity courts against trustees for breaches of fiduciary trust.⁶¹ Interestingly, even one of the drafters of ERISA disagrees with the Court’s interpretation of “equitable relief”, stating that it is “preposterous to think that the ERISA conferees or the ERISA Congress intended to repudiate the law-equity fusion in an ERISA context, and yet would never say a word about it.”⁶²

III. PROBLEMS WITH ERISA AND HEALTH INSURANCE COVERAGE

The majority of courts interpret ERISA to preempt all state remedy laws, both statutory and common law. These courts are incorrect for several reasons: first, Congress’s original intent in passing ERISA; second, the equitable principles of justice in the law; and third, principles of statutory construction. All three reasons support interpreting ERISA to allow at least some state law remedies, particularly those which apply specifically to insurance companies, to apply to certain fiduciary claims in order to protect plan participants and beneficiaries, and ensure that plan fiduciaries exercise the utmost care and diligence in plan decisions, thereby always placing the participants’ needs first.⁶³

A. CASE LAW

Lack of fiduciary loyalty leads to the majority of lawsuits and is a major problem in employee welfare benefit law. Plan fiduciaries must be

⁵⁹ *Mertens v. Hewitt Assoc.*, 508 U.S. 248, 255 (1993) (citing *United States v. Burke*, 504 U.S. 229, 238 (1992)).

⁶⁰ *Id.* at 262-63.

⁶¹ *Id.* at 255-56.

⁶² Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1333 (citing Letter from Michael S. Gordon to John Langbein (June 14, 2002)).

⁶³ It is of note that these duties are what fiduciaries are *supposed* to be held to and follow: the duty of loyalty (to discharge duties *solely* in the interest of participants and beneficiaries, *see* 26 U.S.C. § 404(a)(1)(A) (2006)) and the duty of prudence (to act with the skill and diligence of a prudent person in that position); 26 U.S.C. § 404(a)(1)(B) (2006)).

held to a higher standard, and Congress intended for the courts to hold fiduciaries to this higher standard by incorporating the fiduciaries' duties in section 404(a). Courts have, for the most part, failed to do so (and in some circumstances, specifically excused those fiduciary duties as not applying for HMOs – a real travesty of justice),⁶⁴ and have failed in their administration of justice, as it is highly unlikely that Congress intended to leave participants – the very people sought to be protected – without equitable redress for serious harms committed under ERISA plans.

In *Massachusetts Mutual Life Ins. Co. v. Russell*, the employee-beneficiary, Doris Russell, received her health insurance through an employee benefit plan funded by her employer, Mass. Mutual.⁶⁵ She became disabled due to a back ailment and received benefits for about five months, when the disability committee terminated her benefits based on an orthopedic surgeon's report. She requested an internal review and submitted a report from her psychiatrist indicating that she suffered from a psychosomatic disability, rather than an orthopedic illness. Her benefits were reinstated about five months later, after an examination by another psychiatrist.⁶⁶

She claimed to have been injured from the improper refusal to pay benefits because it forced her disabled husband to cash out his retirement savings, which in turn aggravated the psychological condition that caused her back ailment. Her complaint was based on both state law and on ERISA.⁶⁷ The case was removed to federal court, and the court granted Mass. Mutual's motion for summary judgment, holding that the state law claims were pre-empted by ERISA and the claims for extra-contractual damages and punitive damages were barred under ERISA.⁶⁸ The Ninth Circuit Court of Appeals agreed that the state law claims were preempted, but held that she alleged a cause of action under ERISA, as taking 132 days to process her claim violated a fiduciary's obligation to process claims in

⁶⁴ See *Pegram v. Herdrich*, 530 U.S. 211, 231-35 (2000) (holding that HMOs are not meant to be fiduciaries and held to fiduciary duties under ERISA when making mixed eligibility decisions for participants and beneficiaries; holding the HMOs to such a standard would "in effect . . . be nothing less than elimination of the for-profit HMO" as the court must allow the HMO to make decisions influenced by financial incentives, even if it to the detriment of the participant or beneficiary).

⁶⁵ *Mass. Mut. Life Ins. Co. v. Russell*, 473 U.S. 134, 136 (1985).

⁶⁶ *Id.*

⁶⁷ *Id.* at 137.

⁶⁸ *Id.*

good faith and in a diligent manner.⁶⁹ Thus, the appeals court reasoned, this violation gave rise to a cause of action under § 409(a) that could be asserted by a plan beneficiary pursuant to § 502(a)(2); and under § 409(a), the court has discretion to award “such other equitable or remedial relief as the court may deem appropriate.”⁷⁰ The appeals court believed it had “wide discretion as to the damages to be awarded”, including compensatory and punitive damages.⁷¹ The Ninth Circuit held that punitive damages *are* recoverable under § 409(a) if the fiduciary “acted with actual malice or wanton indifference to the rights of a participant or beneficiary”, and the court believed this result was supported by the text of § 409(a) and the congressional purpose of providing broad remedies to prevent violations of the Act.⁷²

The Supreme Court reversed, holding that under § 409, the remedy of “such other equitable or remedial relief as the court may deem appropriate” is only available to the plan, not to an individual.⁷³ The Court decided that, since Congress’ focus was on protecting mismanagement of pension plans, the intent was to exclude individual recovery to beneficiaries for fiduciary breaches.⁷⁴ The Court stated that a “fair contextual reading of the statute makes it abundantly clear that its draftsmen were primarily concerned with the possible misuse of plan assets, and with remedies that would protect the entire plan, rather than with the rights of an individual beneficiary”, thus eliminating any hope for emotional compensatory damages to the plaintiff, based on a whim of the Court.⁷⁵ The Court candidly and unabashedly wrote off the need to protect the beneficiaries’ interests in a single paragraph.⁷⁶ Conclusively, the court held that plaintiffs may not recover extra-contractual damages, stating it was “reluctant to tamper with an enforcement scheme crafted with such evident care as the

⁶⁹ *Id.* at 137-38.

⁷⁰ *Id.* at 138.

⁷¹ *Russell*, 473 U.S. at 138.

⁷² *Id.*

⁷³ *Id.* at 140.

⁷⁴ *Id.*

⁷⁵ *Id.* at 142.

⁷⁶ *Id.* at 142-43 (the Court states: “It is of course true that the fiduciary obligations of plan administrators are to serve the interest of participants and beneficiaries and, specifically, to provide them with the benefits authorized by the plan. But the principal statutory duties imposed on the trustees relate to the proper management, administration, and investment of fund assets, the maintenance of proper records, the disclosure of specified information, and the avoidance of conflicts of interest.”).

one in ERISA.”⁷⁷ In a subsequent case, the Court noted that this plaintiff had ultimately received all the benefits owed to her, albeit late. Her additional claim for consequential damages was for the delay in processing her claim, and ERISA, the Court concluded, “does not provide a remedy for this type of injury.”⁷⁸

In *Pilot Life*, Dedeaux, the plaintiff-employee, sustained back injuries at work and brought a claim under his employer’s long-term disability plan. The insurer, Pilot Life, processed disability claims and determined who received disability benefits.⁷⁹ Pilot Life initially approved his benefits, and then cancelled them after two years, followed by a three-year period of several benefit reinstatements and terminations by Pilot Life.⁸⁰ Dedeaux brought suit in federal court, citing tortious breach of contract, breach of fiduciary duties, and fraud.⁸¹ He sought damages for mental and emotional distress, along with punitive and exemplary damages, totaling \$750,000.⁸² All of his claims were under state tort law, and not ERISA.⁸³

The district court granted Pilot Life’s motion for summary judgment, ruling that all of the claims were preempted under ERISA.⁸⁴ The Fifth Circuit reversed, holding that the law was saved under the savings clause because it affected the “relationship between the insurer and the insured,” thus placing it within the ‘business of insurance’ under the McCarran-Ferguson Act, and therefore qualifying as a law which regulates insurance.⁸⁵

The Supreme Court reversed. First, it noted that Dedeaux’s claims clearly ‘related to’ the ERISA plan were preempted under section 514(a).⁸⁶ Second, the Court rejected that the Mississippi law of bad faith is saved by the savings clause, because it is a general tort law that did not specifically

⁷⁷ *Russell*, 473 U.S. at 147.

⁷⁸ *LaRue v. DeWolff, Boberg & Assoc.*, 552 U.S. 248, 254 (2008).

⁷⁹ *Pilot Life Ins. Co. v. Dedeaux*, 481 U.S. 41, 43 (1987).

⁸⁰ *Id.*

⁸¹ *Id.*

⁸² *Id.* at 43-44.

⁸³ *Id.*

⁸⁴ *Id.* at 44.

⁸⁵ *Dedeaux v. Pilot Life Ins. Co.*, 770 F.2d 1311, 1316 (5th Cir. 1985) (citing *Metro. Life Ins. Co. v. Mass.*, 105 S. Ct. 2380, 2391 (1985)).

⁸⁶ *Pilot Life Ins. Co.*, 481 U.S. at 47-48.

apply to insurance companies, and thus did not regulate insurance.⁸⁷ The Court distinguished this case from *Metropolitan Life Ins.*, stating:

Unlike the mandated-benefits law at issue in *Metropolitan Life*, the Mississippi common law of bad faith does not effect a spreading of policyholder risk. The state common law of bad faith may be said to concern “the policy relationship between the insurer and the insured.” The connection to the insurer-insured relationship is attenuated at best, however. In contrast to the mandated-benefits law in *Metropolitan Life*, the common law of bad faith does not define the terms of the relationship between the insurer and the insured; it declares only that, whatever terms have been agreed upon in the insurance contract, a breach of that contract may in certain circumstances allow the policyholder to obtain punitive damages. The state common law of bad faith is therefore no more “integral” to the insurer-insured relationship than any State’s general contract law is integral to a contract made in that State. Finally, as we have just noted, Mississippi’s law of bad faith, even if associated with the insurance industry, has developed from general principles of tort and contract law available in any Mississippi breach of contract case.⁸⁸

The Court held that in order to be saved under the savings clause the state law must specifically regulate insurance.

In *Varity Corp. v. Howe*, the employer, Varity Corporation, owned a subsidiary, Massey-Ferguson (“MF”), which employed the plaintiffs and provided a self-funded employee welfare plan.⁸⁹ Varity determined that

⁸⁷ *Id.* at 50 (“A common-sense view of the word ‘regulates’ would lead to the conclusion that in order to regulate insurance, a law must not just have an impact on the insurance industry, but must be specifically directed toward that industry. Even though the Mississippi Supreme Court has identified its law of bad faith with the insurance industry, the roots of this law are firmly planted in the general principles of Mississippi tort and contract law. Any breach of contract, and not merely breach of an insurance contract, may lead to liability for punitive damages under Mississippi law.”).

⁸⁸ *Id.* at 50-51.

⁸⁹ *Varity Corp. v. Howe*, 516 U.S. 489, 492 (1996). (Note that under the ERISA deemer clause, a self-funded plan is exempt from compliance with state insurance regulations and statutes.)

several MF divisions were losing too much money, and concocted a plan to place all of the unstable divisions and debt, and the employees, into another subsidiary, Massey Combines (MC’).⁹⁰ The makers of the plan acknowledged the possibility that MC would fail, but saw that outcome as a victory because it would eliminate the unstable divisions and the debt transferred to MC.⁹¹ If Varsity did not form the separately incorporated subsidiary, then Varsity itself would have to pay for the debt and the ERISA plan benefits for the unstable divisions.⁹² Instead of terminating the plan, which would have likely resulted in the employees leaving to find new employment (and “voluntarily release[ing] Massey-Ferguson from its obligation to provide them benefits”), Varsity essentially made MC the new employer and switched the employees to a new plan that was funded by MC.⁹³ However, the employees had to elect to switch employers and thus to switch plans. To persuade them, Varsity held a meeting and presented the plan, passed out documents which represented that the benefits would remain the same and were safe, but noted that employment conditions in the future depended on MC’s success.⁹⁴ Unfortunately, all the employees agreed to Varsity’s plan, and Varsity also took the liberty of assigning the benefit obligations to 4,000 retired workers to MC.⁹⁵

The lower court found that MC was insolvent from its very first day. It ended its first year with an \$88 million loss and its second year in receivership, thus terminating the employees’ non-pension benefits.⁹⁶ The Supreme Court held that Varsity was a fiduciary of the plan, as it was both the employer and administrator of the plan and was acting in a fiduciary capacity during the meeting with employees.⁹⁷ The specific context of the events supported that Varsity was exercising discretionary authority on the plan’s management or administration when it made the benefit representations to the employees.⁹⁸ The main message represented to the employees at the meeting, by designated fiduciaries in the plan documents, was that transferring to MC would not undermine their benefits. This constituted conveying benefit information to participants.⁹⁹ Such

⁹⁰ *Id.* at 492-93.

⁹¹ *See id.* at 493.

⁹² *Id.*

⁹³ *See id.* at 493-94.

⁹⁴ *Id.* at 494, 500-01.

⁹⁵ *Varsity Corp.*, 516 U.S. at 494.

⁹⁶ *Id.*

⁹⁷ *Id.* at 498.

⁹⁸ *Id.* at 498-99.

⁹⁹ *Id.* at 502-03.

information is specifically required under ERISA to allow participants the ability to make an informed choice about continued participation in a new benefit plan.¹⁰⁰ The district court concluded that since the fiduciaries knew that there was a high likelihood that the employees' plans were not safe and would not remain the same, their statements were materially misleading.¹⁰¹ The Supreme Court held that knowingly deceiving plan participants and beneficiaries violates the duty of loyalty: "lying is inconsistent with the duty of loyalty owed by all fiduciaries and codified in section 404(a)(1) of ERISA."¹⁰² The Court further noted that trust law requires trustees to deal "fairly and honestly with beneficiaries."¹⁰³

In determining the relief warranted under ERISA, the Court held that individual relief was appropriate under section 502(a)(3), which provides that a participant, beneficiary, or fiduciary may bring suit "(A) to enjoin any act or practice which violates any provision of this title or the terms of the plan, or (B) to obtain other appropriate equitable relief (i) to redress such violations or (ii) to enforce any provisions of this title or the terms of the plan."¹⁰⁴ The Court stated that individual relief is authorized under ERISA section 502(l) for violations of section 502(a)(5), and that section 502(a)(3) is identical to 502(a)(5), except that it allows suit to be brought by the Secretary of Labor.¹⁰⁵ Further, the legislative history supported a broad reading of section 502 to permit broad remedies for redressing or preventing violations of ERISA.¹⁰⁶ The Court further noted that it would be "hard to imagine why Congress would want to immunize breaches of fiduciary obligation that harm individuals by denying injured

¹⁰⁰ *Id.* at 502, 505 (citing ERISA §§ 102, 104(b)(1), 105(a)). The Court rejected that these statements were made as the employer in a plan sponsor (thus in a manner similar to amending a plan, which is not a fiduciary act) because the statements were about the future of benefits, which is something that plan administrators regularly communicate to participants and beneficiaries. *Id.* at 505.

¹⁰¹ *Varity Corp.*, 516 U.S. at 505.

¹⁰² *Id.* (quoting *Peoria Union Stock Yards Co. v. Penn. Mut. Life Ins. Co.*, 698 F.2d 320, 326 (7th Cir. 1983)).

¹⁰³ *Id.* (citing GEORGE GLEASON BOGERT & GEORGE TAYLOR BOGERT, *THE LAW OF TRUSTS AND TRUSTEES* § 543, at 218-19 (2d. ed. 1993); 2A AUSTIN WAKEMAN SCOTT & WILLIAM FRANKLIN FRATCHER, *THE LAW OF TRUSTS* § 170, at 311-12 (4th ed. 1987)).

¹⁰⁴ 29 U.S.C. § 1132(a) (2006).

¹⁰⁵ *Varity Corp.*, 516 U.S. at 510.

¹⁰⁶ *Id.* at 512 (citing S. REP. NO. 93-127, at 35 (1973), *reprinted in* 1974 U.S.C.C.A.N. 4838).

beneficiaries a remedy.”¹⁰⁷ Therefore, individual relief is allowed under ERISA for fiduciary breaches.

In *Andrews-Clarke v. Travelers Insurance*, arguably one of the most¹⁰⁸ egregious abuses of fiduciary conduct, the plaintiff obtained health insurance through her employer, AT&T, and her children and husband were beneficiaries of the policy.¹⁰⁹ Her husband had a drinking problem and was admitted to a hospital for alcohol detoxification and medical evaluation. The hospital notified Greenspring, the review provider who, under the plan, must pre-approve treatment. Greenspring refused to approve the treatment even though the insurance policy specifically entitled the insured and beneficiaries to “at least one thirty-day inpatient rehabilitation program per year.”¹¹⁰ Greenspring approved only a five-day stay, and he was discharged after five days, “with a diagnosis of alcohol dependence, alcohol withdrawal symptoms [and] elevated liver function.”¹¹¹ Twenty-five days later, he resumed drinking and admitted himself to another hospital. Despite the terms of the insurance plan, Greenspring approved only an eight-day stay.¹¹² After being discharged, he consumed a large amount of alcohol, cocaine, prescription drugs, and attempted to commit suicide. After a commitment hearing, the court clinic requested Greenspring’s approval for a thirty-day treatment at a private hospital, which it declined. Mr. Clarke was committed to Southeastern Correctional Center at Bridgewater for his detoxification and rehabilitation, where he received “little in the way of therapy or treatment” and was forcibly raped by a fellow inmate.¹¹³ Upon release, he resumed drinking and committed suicide.¹¹⁴

Mrs. Andrews-Clarke brought suit against Travelers and Greenspring, claiming Clarke’s death was the direct and foreseeable result of the improper refusal of Travelers and its agent Greenspring to authorize appropriate medical and psychiatric treatment during Clarke’s repeated hospitalizations for alcoholism in 1994. Her claims included breach of contract, medical malpractice, wrongful death, loss of parental and spousal

¹⁰⁷ *Id.* at 513.

¹⁰⁸ Although certainly not *the most* egregious. *See infra* Part IV.

¹⁰⁹ *Andrews-Clarke v. Travelers Ins. Co.*, 984 F. Supp. 49, 50 (D. Mass. 1997).

¹¹⁰ *Id.* at 50-51.

¹¹¹ *Id.* at 51.

¹¹² *Id.*

¹¹³ *Id.*

¹¹⁴ *Id.* at 52.

consortium, intentional and negligent infliction of emotional distress, and specific violations of the Massachusetts consumer protection laws.¹¹⁵

The district court spoke on the importance of breach of contract claims, noting that such causes of action “pre-date [the] Magna Carta” and “[are] the very bedrock of our notion of individual autonomy and property rights”, arguing that “[o]ur entire capitalist structure depends on it.”¹¹⁶ Clearly regretting what it understood the law to be, the court granted Travelers’ motion to dismiss, effectively “slam[ming] the courthouse doors in [Andrews-Clarke’s] face and leav[ing] her without any remedy.”¹¹⁷ The court noted that it was just another example of “the glaring need for Congress to amend ERISA to account for the changing realities of the modern health care system” because ERISA had “evolved into a shield of immunity that protects health insurers, utilization review providers, and other managed care entities from potential liability for the consequences of their wrongful denial of health benefits.”¹¹⁸ The judge acknowledged that the plaintiff’s claims would be cognizable if not controlled by ERISA. The federal judge concluded in a candid statement:

Employee health benefit plans lack security because of the *de facto* immunity that the law now confers upon insurers and utilization review providers associated with such plans. Unfortunately, to date, “ERISA [has proven an excellent example of the classic observation that it is a

¹¹⁵ *Andrews-Clark*, 984 F. Supp. at 52.

¹¹⁶ *Id.* at 52-53 (citing E. ALLAN FARNSWORTH, CONTRACTS §§ 1.4-1.6 (2d ed. 1990)).

¹¹⁷ *Id.* at 53.

¹¹⁸ *Id.* (citing *Tolton v. Am. Biodyne, Inc.*, 48 F.3d 937, 943 (6th Cir. 1995) (“One consequence of ERISA preemption, therefore, is that plan beneficiaries or participants bringing certain types of state actions - such as wrongful death - may be left without a meaningful remedy.”); *Corcoran v. United HealthCare, Inc.*, 965 F.2d 1321, 1338 (5th Cir. 1992) (“The result ERISA compels us to reach means the [Plaintiff has] no remedy, state or federal, for what may have been a serious mistake.”); *Turner v. Fallon Cmty. Health Plan Inc.*, 953 F. Supp. 419, 424 (D. Mass. 1997) (Gorton, J.) (“An unfortunate consequence of ERISA preemption is, therefore, that plan beneficiaries or participants who bring certain kinds of state actions, e.g., wrongful death, may be left without a meaningful remedy. . . . Sadly, the case at bar compels a like result. Plaintiff’s state common law claims are preempted by the broadly sweeping arm of ERISA. Plaintiff is left without any meaningful remedy even if he were to establish that [the insurer] wrongfully refused to provide the [bone marrow transplant] his wife urgently sought.”), *aff’d*, 127 F.3d 196 (1st Cir. 1997)).

great deal more difficult for Congress to correct flawed statutes than it is to enact them in the first place . . . because interests coalesce around the advantageous aspects of the status quo. Although the alleged conduct [of wrongfully denying benefits that were clearly due to the beneficiary, resulting in his death] of Travelers and Greenspring in this case is extraordinarily troubling, even more disturbing to this Court is the failure of Congress to amend a statute that, due to the changing realities of the modern health care system, has gone conspicuously awry from its original intent.

Does anyone care? Do you?¹¹⁹

This judge brings to the forefront the serious concerns implicit in ERISA law – the lack of equitable relief allowed to injured beneficiaries in these circumstances.

In *Aetna Health, Inc. v. Davila*, there were two separate cases of egregious HMO behavior, consolidated because the preemption issues were the same for both.¹²⁰ Juan Davila, covered under his employer's benefit plan, was prescribed Vioxx for arthritis pain by his treating doctor.¹²¹ Aetna, the plan administrator that reviews requests for coverage and pays providers, declined the prescription and advised that it would only cover Naprosyn; Davila ingested the Naprosyn and suffered a severe reaction that required hospitalization and extensive treatment.¹²² Ruby Calad, also covered under an ERISA plan, underwent surgery and her doctor recommended an extended hospital stay to prevent post-surgery complications.¹²³ Her plan administrator's discharge nurse, a CIGNA employee, concluded that Calad did not need the extended stay and advised CIGNA to deny coverage, which it did.¹²⁴ Forced to return home too early, she suffered post-surgery complications and required hospitalization.¹²⁵

¹¹⁹ *Id.* at 64-65 (first alteration in original) (footnote omitted) (quoting Catherine L. Fist, *The Last Article About the Language of ERISA Preemption?: A Case Study of the Failure of Textualism*, 33 HARV. J. ON LEGIS. 35, 99 (1996)).

¹²⁰ *Aetna Health, Inc. v. Davila*, 542 U.S. 200, 204 (2004).

¹²¹ *Id.* at 205.

¹²² *Id.*

¹²³ *Id.*

¹²⁴ *Id.*

¹²⁵ *Id.*

Both plaintiffs sued the plan administrators, alleging violations of a Texas bad-faith insurance statute.¹²⁶ The district court held that ERISA preempted the claims, and, as the plaintiffs refused to amend their complaints to allege ERISA claims, dismissed the complaints with prejudice.¹²⁷ On appeal, the Fifth Circuit held that plaintiffs' claims were cognizable actions under ERISA's "[section] 502(a)(1)(B), which provides a cause of action for the recovery of wrongfully denied benefits, and [section] 502(a)(2), which allows suit against a plan fiduciary for breaches of fiduciary duty to the plan."¹²⁸ But because the decisions were mixed eligibility and treatment decisions by HMOs, they were not fiduciary in nature, and no relief was available under section 502(a)(1).¹²⁹

The Supreme Court noted that the plaintiffs only complained about the denials of coverage promised under the terms of the plan, and that they were fully capable of, upon denial, paying for the treatment themselves and then seeking reimbursement through a section 502(a)(1)(B) action, or through a preliminary injunction.¹³⁰ The Court held that essentially the claims were for the wrongful denial of benefits, which was a claim available under section 502, and that when a state law cause of action duplicates, supplements, or supplants the ERISA civil enforcement remedy, it conflicts with the clear congressional intent to make the ERISA remedy exclusive and therefore is preempted.¹³¹ The Court rejected plaintiffs' argument that its claim was saved under the savings clause, which specifically regulates insurance, instead holding that because, as held in *Pilot Life Ins. Co.*, Congress set out a comprehensive remedial scheme in ERISA, and "[t]he policy choices reflected in the inclusion of certain remedies and the exclusion of others under the federal scheme would be completely undermined if ERISA-plan participants and beneficiaries were

¹²⁶ *Davila*, 542 U.S. at 205 (citing TEX. CIV. PRAC. & REM. § 88.001 (West 2011)).

¹²⁷ *Id.*

¹²⁸ *Id.* at 206.

¹²⁹ *Id.* (citing *Pegram v. Hedrich*, 530 U.S. 211 (2000)).

¹³⁰ *Id.* at 211. Clearly the Court (as it was a unanimous decision) did not understand the position the plaintiffs were in and were unable to empathize or sympathize: what person can afford to pay for a hospital visit on his own, without insurance? How is this a reasonable suggestion? I cannot imagine the brazenness in even suggesting that the plaintiff take on this responsibility, when his insurance should be covering it, and failed to do so out of severe negligence. And the Court appears to support such an act, or to shield the fiduciary, when justice and Congressional intent clearly foresee a different outcome.

¹³¹ *Id.* at 208-09, 213-14.

free to obtain remedies under state law that Congress rejected in ERISA.”¹³²

Justice Ginsburg’s concurrence is enlightening; she stated that she joined the “rising judicial chorus urging that Congress and [this] Court revisit what is an unjust and increasingly tangled ERISA regime.”¹³³ She noted that the Court’s interpretations of preemption and the “equitable relief” clause under section 502(a)(3) had left a “regulatory vacuum” because “virtually all state law remedies are preempted but very few federal substitutes are provided.”¹³⁴ She went on to give specific examples of situations in which fiduciary breaches have left the beneficiary in a deficient position due to an inability to receive ‘make-whole’ relief under ERISA.¹³⁵ Thus, we are left with a bad taste of ‘justice’ under ERISA: that some members of the Court recognize that it is misinterpreting or incorrectly applying ERISA, leaving plaintiffs harmed yet with little or no relief.

In *Metropolitan Life Ins. Co. v. Glenn*, the plaintiff was diagnosed with a “severe dilated cardiomyopathy, a heart condition”, and she applied for long term disability under her employer’s employee welfare benefit plan.¹³⁶ MetLife, the insurance provider on the plan, approved her for the initial 24 months of benefits, concluding that she could not perform the material duties of her job.¹³⁷ MetLife directed her to a law firm which would help her apply for Social Security disability benefits, and “an

¹³² *Davila*, 542 U.S. at 217 (alteration in original) (quoting *Pilot Life Ins. Co. v. Dedeaux*, 481 U.S. 41, 54 (1987)) (internal quotation marks omitted).

¹³³ *Id.* at 222 (Ginsburg, J., concurring) (quoting *DiFelice v. Aetna U.S. Healthcare*, 346 F.3d 442, 453 (3d Cir. 2003) (Becker, J., concurring)).

¹³⁴ *Id.* (quoting *DeFilice*, 346 F.3d at 456-57 (Becker, J., concurring)).

¹³⁵ *Id.* at 222-23 (citing *Mass. Mut. Life Ins. Co. v. Russell*, 473 U.S. 134, 148 (1985) (“[T]he Court stated, in dicta: [T]here is a stark absence - in [ERISA] itself and in its legislative history - of any reference to an intention to authorize the recovery of extracontractual damages for consequential injuries.”) (second and third alterations in original) (internal quotation marks omitted); *Mertens v. Hewitt Assoc.’s*, 508 U.S. 248, 255-56 (1993) (“[T]he Court held that [section] 502(a)(3)’s term ‘equitable relief’ . . . refer[s] to those categories of relief that were typically available in equity (such as injunction, mandamus, and restitution, but not compensatory damages).”) (third alteration in original) (internal quotation marks omitted); *Great-West Life & Annuity Ins. Co. v. Knudson*, 534 U.S. 204, 221 (2002) (“the Court ruled that, as [section] 502(a)(3), by its terms, only allows for equitable relief, the provision excludes “the imposition of personal liability ... for a contractual obligation to pay money.”) (internal quotation marks omitted)).

¹³⁶ *Metro. Life Ins. v. Glenn*, 128 S. Ct. 2343, 2346 (2008).

¹³⁷ *Id.*

Administrative Law Judge found that her illness prevented her not only from performing her own job but also from performing any jobs [for which she could qualify] existing in significant numbers in the national economy,” thus meeting the standard for Social Security benefits.¹³⁸ She was granted permanent disability benefits, the majority of which went to MetLife and the rest of which went to the lawyers.¹³⁹ MetLife subsequently denied her benefits beyond the 24-month mark, determining that she was “capable of performing full time sedentary work,” applying a standard similar to the one used by the administrative judge, who had found her incapable of such work.¹⁴⁰

She brought suit for judicial review of this denial of benefits, and the district court granted MetLife’s motion for judgment because the plan granted the plan administrator discretionary authority in deciding benefits.¹⁴¹ The Sixth Circuit set aside the denial of benefits because of:

(1) the conflict of interest [arising from MetLife’s authority to determine whether it was obligated to pay benefits to an employee]; (2) MetLife’s failure to reconcile its own conclusion that Glenn could work in other jobs with the Social Security Administration’s conclusion that she could not; (3) MetLife’s focus upon one treating physician report suggesting that Glenn could work in other jobs at the expense of other, more detailed treating physician reports indicating that she could not; (4) MetLife’s failure to provide all of the treating physician reports to its own hired experts; and (5) MetLife’s failure to take account of evidence indicating that stress aggravated Glenn’s condition.¹⁴²

The Supreme Court affirmed, holding that when a plan administrator both evaluates claims for benefits and pays benefits on claims, it creates a conflict of interest that a court may review for abuse of discretion in denying benefits, because “every dollar provided in benefits is a dollar spent by ... the employer; and every dollar saved . . . is a dollar in

¹³⁸ *Id.* (internal quotation marks omitted).

¹³⁹ *Id.* at 2346-47.

¹⁴⁰ *Id.* at 2347.

¹⁴¹ *Glenn v. MetLife*, 461 F.3d 660, 665 (6th Cir. 2006).

¹⁴² *Metro. Life Ins.*, 128 S. Ct. at 2347.

[the employer's] pocket.”¹⁴³ A troubling example appears when “[t]he employer's fiduciary interest may counsel in favor of granting a borderline claim while its immediate financial interest counsels to the contrary”, and “the employer has an interest . . . conflicting with that of the beneficiaries.”¹⁴⁴ Here, the Court affirmed the lower court's review of the benefit denial, and in doing so, took a step forward in beneficiary protection.

B. GOOD MORNING AMERICA INVESTIGATES UNFAIR CLAIMS PRACTICES IN DISABILITY INSURANCE

In 2008, Good Morning America, a daily news program aired by ABC, broke several investigative news stories about disability insurers and unfair claims practices. In April, the story of Susan Kristoff displayed to the world the dishonest practices and lengths one insurer was willing to go to avoid paying her disability claim. She was diagnosed with Stage IV metastatic breast cancer and had several doctors' accounts confirming that she was disabled.¹⁴⁵ The contract with Cigna, the disability insurer, provided that she be paid 60% of pre-disability income upon disability.¹⁴⁶ But Cigna denied her claims, and she spent two years attempting to furnish the “additional information” constantly requested by Cigna as “necessary” to further review her claim.¹⁴⁷ Finally, she hired a lawyer and wrote to Good Morning America about her plight. Good Morning America contacted Cigna to attempt to resolve the issue.¹⁴⁸ In response, Cigna sent a

¹⁴³ *Id.* at 2348 (quoting *Bruch v. Firestone Tire & Rubber Co.*, 828 F.2d 134, 144 (3d Cir. 1987)).

¹⁴⁴ *Id.* (internal quotation marks omitted) (citing RESTATEMENT (SECOND) OF TRUSTS § 187 cmt. d (1959)); *see also* *Firestone Tire & Rubber Co. v. Bruch*, 489 U.S. at 115 (1989) (citing that Restatement comment); *cf.* BLACK'S LAW DICTIONARY 319 (8th ed. 2004) (A conflict of interest is “[a] real or seeming incompatibility between one's private interests and one's public or fiduciary duties.”).

¹⁴⁵ *See* Chris Cuomo & Gerry Wagschal, *Denied: Fighting for Insurance Coverage*, ABC NEWS (Jun. 27, 2008), <http://abcnews.go.com/GMA/story?id=5257491&page=1>; Brent Adams, *Every Bad Faith Insurance Victim Cannot Have Good Morning America As Their Lawyer*, INJURYBOARD BLOG FAYETTEVILLE, N.C. (Feb. 1, 2009 9:20 PM), <http://fayetteville.injuryboard.com/miscellaneous/every-badfaith-insurance-victim-cannot-have.aspx?googleid=256444>.

¹⁴⁶ Adams, *supra* note 145.

¹⁴⁷ Cuomo & Wagschal, *supra* note 145; Adams, *supra* note 145.

¹⁴⁸ Cuomo & Wagschal, *supra* note 145; Adams, *supra* note 145.

canned statement to Good Morning America and promptly approved Kristoff's claim, denying that it was because of the show's involvement.¹⁴⁹

Another story published online in 2009 described Charles Tucker's fight with Standard Insurance over disability benefits. Tucker, a 48 year-old accountant, suffered from a severe and debilitating form of multiple sclerosis.¹⁵⁰ He had been paying for long-term disability insurance coverage, but when he became too sick to work, he stopped working and filed a claim. He received constant notices requesting more information and for months could not receive a final determination from the insurer.¹⁵¹ Eleven doctors examined Tucker, and all concluded that he had MS. However, the insurer's doctor, without examining Tucker or contacting him about the inquiry, determined that there was insufficient evidence for such a conclusion. Tucker hired an attorney and contacted Good Morning America. The show's anchor contacted the insurer and spoke with a spokeswoman for the company, Susan Pisano, also a lobbyist for America's Health Insurance Plans. The next day, Standard Insurance approved Tucker's claim, but denied that it was because of Good Morning America's investigation.¹⁵²

One personal injury attorney blogged about this seeming epidemic of disability claims denials, arguing that disability claimants are in the worst position to fight denials.¹⁵³ He stated what should be obvious: "The disabled person is the least likely to be able to afford an attorney [because] a major source of income has been taken away."¹⁵⁴ Perhaps his assertion is correct, that most insurers are betting on the fact that denials will *not* be appealed due to lack of money, or fear that the insurer will insist that a fraud is being committed, or lack of understanding of the denial-appeal system.¹⁵⁵ If this is true, then the argument for strict regulatory oversight becomes absolutely necessary, as this will ensure that such claimants will

¹⁴⁹ Cuomo & Wagschal, *supra* note 145; Adams, *supra* note 145.

¹⁵⁰ See Chris Cuomo & Gerry Wagschal, 'GMA' Gets Answers: Man with MS Fights for Long-Term Disability Insurance, ABC NEWS (Feb. 26, 2009), <http://abcnews.go.com/GMA/TheLaw/gma-answers-man-ms-fights-long-term-disability/story?id=6689847>.

¹⁵¹ *Id.*

¹⁵² *Id.*

¹⁵³ See Frank J. Dito Jr, *Good Morning America Airs Report about Disability Insurance*, NEW YORK INJURY LAW BLOG (Nov. 19, 2009), <http://nyinjurylawupdateblog.com/good-morning-america-air-report-about-disability-insurance>.

¹⁵⁴ *Id.*

¹⁵⁵ *Id.*

not only have fair reviews of their claims but have recourse and knowledge of the appropriate recourse.

C. VIEWS AND RECOMMENDATIONS TO RESOLVE ERISA'S REMEDIES PROBLEM

Several authors write of different approaches for allowing extra-contractual relief under ERISA. Three authors argue that compensatory (i.e., money) damages are appropriate under ERISA's "other appropriate equitable relief" clause in section 502, because such relief is allowed in trust law.¹⁵⁶ Another argues that plaintiffs use RICO claims to receive appropriate relief.¹⁵⁷ Yet another approach argues that state insurance unfair practices statutes must be saved from preemption, as section 514(b)(2)(B) expressly enforces state insurance laws, while section 502 only generally provides remedies and does not expressly preempt insurance laws.¹⁵⁸

1. Trust Law As A Basis For Additional Remedies To Beneficiaries

In a poignant review of the U.S. Supreme Court's decision in *Aetna v. Davila*, Charlotte Johnson advocates that the Court has not enforced ERISA as it was intended by Congress, and has effectively, but inadvertently, "painted itself into a corner by restrictively interpreting ERISA to preclude compensatory relief to victims of HMO patient treatment decisions."¹⁵⁹ She argues that, as ERISA was enacted before the surge of HMOs, Congress could not have anticipated ERISA's effect and regulation of HMO liability.¹⁶⁰ The Court has time and again decided that, under ERISA, injured participants and beneficiaries may only receive traditional equitable relief, i.e., injunction or restitution, and not compensatory relief, e.g., money damages; thus, the Court has provided a shield for HMOs against plan participant claims.¹⁶¹

¹⁵⁶ Johnson, *supra* note 6, at 1589; Harthill, *supra* note 6, at 722; Langbein, *Trust Law As Regulatory Law*, *supra* note 6, at 1340-41; Langbein, *What ERISA Means by "Equitable"*, *supra* note 6, 1338, 1362.

¹⁵⁷ Lalena J. Turchi, *Health Insurance: Paying the Premium or Paying the Price? – ERISA Preemption and RICO's Recourse*, 5 RUTGERS J. L. & PUB. POL'Y 526, 551-59 (2008).

¹⁵⁸ Bogan, *supra* note 6, at 113-14.

¹⁵⁹ Johnson, *supra* note 6, at 1589-90.

¹⁶⁰ *Id.* at 1590.

¹⁶¹ *Id.*

The only way of effectively getting around this shield was pointed out by Justice Ginsburg in her *Davila* dissent: allow make-whole compensatory relief under section 502(a)(3).¹⁶² The Court's interpretation of "appropriate equitable relief" is strange, in that it refuses to impose personal liability on the defendant, as that would transfer the equitable restitution to one of legal restitution.¹⁶³ However, section 409 *allows* a fiduciary to be personally liable for breaches, thus such personal liability is, in fact, permitted and authorized under ERISA.¹⁶⁴ Johnson also advocates Justice Stevens' position, that Congress intended ERISA to provide a broad framework under which the courts may apply make-whole compensatory relief, due to ERISA's skeletal framework incorporating some facets of trust law combined with Congressional intent to "protect the interests of participants in employee benefit plans."¹⁶⁵ Further, monetary (compensatory) damages are authorized under ERISA because, as shown in *Varity*, section 502(a)(5) permits payment of civil penalties for fiduciary breaches to participants and beneficiaries, and sections 502(a)(5) and 502(a)(3) are nearly identical.¹⁶⁶ Further, Johnson argues that trust law permits monetary equitable relief to individuals, yet the Court's decision to overlook this principle of trust law has resulted in fiduciary breaches, in the HMO context, to not be a breach at all.¹⁶⁷

Johnson concedes that the remedial scheme in ERISA is properly fixed by Congress, not "creativity in the courts."¹⁶⁸ This article is encouraging in its analysis of trust law, however, the concession that only Congress can fix the problem is frustrating. Courts interpret laws, and if they misinterpret the laws, Congress may react and amend the law. Unfortunately, this does not always happen, and the health insurance industry likely has a stronghold on many politicians that prevented such an amendment from being passed.

Professor John Langbein argues that Congress only referenced trust law as a regulatory structure, by using the tenets of loyalty and prudence, but left ERISA skeletal in form to be refined by the judiciary.¹⁶⁹ The Court itself acknowledged that Congress intended the judiciary to look to settled

¹⁶² *Id.* at 1591.

¹⁶³ *Knudson*, 534 U.S. 204, 214 (2002).

¹⁶⁴ *See* 29 U.S.C. § 1109 (2006).

¹⁶⁵ Johnson, *supra* note 6, at 1612 (citing 29 U.S.C. § 1001(b) (2006); *Varity Corp. v. Howe*, 516 U.S. 489, 496-97, 502-03 (2006)).

¹⁶⁶ *Id.* at 1613 (citing *Varity*, 516 U.S. at 510).

¹⁶⁷ *Id.* at 1622-23.

¹⁶⁸ *Id.* at 1617.

¹⁶⁹ Langbein, *What ERISA Means by "Equitable"*, *supra* note 6, at 1326.

common law in shaping ERISA, thus creating a “federal common law of rights and obligations under ERISA-regulated plans.”¹⁷⁰ Langbein illustrates several instances where the Court failed to correctly apply trust law principles, such as in *Russell* by rejecting the plaintiff’s claim for delay damages.¹⁷¹ Delay in making a distribution from a trust fund “has long been understood to be a breach of trust.”¹⁷² In the Uniform Trust Code of 2000, money damages for breaches of trust are included in the available remedies.¹⁷³ In Bogert’s treatise on trust law, breaching trustees “may be directed by the court to pay damages to the beneficiary” and in cases of negligence or misconduct, the beneficiary may have a claim “to recover money damages from the trustee.”¹⁷⁴ Thus, it seems clear that trust law, in fact, does allow beneficiaries to recover compensatory and consequential damages.

In drafting ERISA, Congress surpassed trust law verbiage and further subjected the fiduciary to “such other equitable or remedial relief as the court may deem appropriate,” on top of making the fiduciary liable to the plaintiff for recovery of the loss incurred, profits made by the fiduciary in the breach, and any gains made from the breach.¹⁷⁵ This phrase is also found in section 502(a)(3), and is more expansive than the trust law standard, and it is the belief of several scholars that this language is meant to provide compensatory and consequential relief, as Congress was simply wording it in a way to address fairness.¹⁷⁶ If the Court refuses to utilize trust law principles, then perhaps the statutory construction should persuade it that, because this phrase surpasses trust law verbiage, Congress intended for fairness to govern, thus allowing additional recovery for egregious breaches.

¹⁷⁰ *Firestone Tire & Rubber Co. v. Bruch*, 489 U.S. 101, 110-11 (1989).

¹⁷¹ *Mass. Mut. Life Ins. Co. v. Russell*, 473 U.S. 134, 144 (1985).

¹⁷² Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1328.

¹⁷³ *Id.* at 1336 (citing UNIF. TRUST CODE §§ 106, 1002 (amended 2001)).

¹⁷⁴ Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1337 (citing GEORGE GLEASON BOGERT & GEORGE TAYLOR BOGERT, *THE LAW OF TRUSTS AND TRUSTEES* § 862, at 34, 38-39 (rev. 2d ed. 1982)).

¹⁷⁵ See ERISA § 409(a), 29 U.S.C. § 1109 (1988); Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1335.

¹⁷⁶ Langbein, *What ERISA Means by “Equitable”*, *supra* note 6, at 1335-36.

2. Applying the Savings Clause with Force: State Unfair Insurance Practices Statutes Should be Saved from Preemption under the Pilot Life Rationale

In an article written before *Davila*, Professor Donald Bogan argues that state unfair insurance practices laws should be saved from preemption under the savings clause.¹⁷⁷ These laws are specifically aimed at the regulation of the insurance industry, fulfill the requirements in the savings clause, and are saved from preemption under section 502 because of the express language in section 514(b)(2)(B), authorizing and enforcing state insurance laws to ERISA plans.¹⁷⁸ In analyzing *Pilot Life*, he argues that the Court only preempted the law at issue because it was a general bad faith law that did not specifically regulate insurance, and as such was preempted by the remedies in section 502.¹⁷⁹ In this case the Court made a landmark decision: that section 502 was intended by Congress as “the exclusive vehicle for actions by ERISA-plan participants . . . asserting improper processing of a claim for benefits.”¹⁸⁰

Bogan points out that state unfair insurance practices statutes do more than simply provide punitive damages remedies, they “establish and define a standard of care owed by the insurer to the insured that attaches to every insurance policy.”¹⁸¹ The Supreme Court has, in the past, declared that state insurance laws affect an integral part of the policy relationship between the insurer and insured, and are saved from preemption.¹⁸² Bogan

¹⁷⁷ Bogan, *supra* note 6, at 113-14.

¹⁷⁸ *Id.* at 114.

¹⁷⁹ *Id.* at 124-25 (citing *Pilot Life Ins. Co. v. Dedeaux*, 481 U.S. 41, 50 (1987)).

¹⁸⁰ *Id.* at 126 (citing *Pilot Life*, 481 U.S. at 52; *Mass. Mut. Life Ins. Co. v. Russell*, 473 U.S. 134, 147 (1985)). In a subsequent decision by the Eleventh Circuit Court of Appeals, a claim under the Florida unfair insurance practices statute for disability benefits was denied. It was denied for two reasons: while it directly regulated insurance, it failed to meet all three prongs under the McCarran Ferguson test, thus precluding it from regulating the “business of insurance.” The court concluded that this plus the *Pilot Life* analysis forced preemption due to section 502. *Id.* at 131-32 (citing *Anschultz v. Conn. Gen. Life Ins. Co.*, 850 F.2d 1467, 1468-69 (11th Cir. 1988)).

¹⁸¹ *Id.* at 133.

¹⁸² *See, e.g., Metro. Life Ins. Co. v. Mass.*, 471 U.S. 724, 743-44 (1985); Bogan, *supra* note 6, at 134 (citing *UNUM Life Ins. Co. v. Ward*, 526 U.S. 358, 374 (1999) (holding that a California notice-prejudice law, requiring the insurer to establish prejudice before denying a claim filed late, was saved from preemption because it served as an integral part of the policy relationship)).

asks: why are state insurance unfair practice statutes not also saved, when they clearly affect every insurance contract and “effectively create mandatory contract terms that require insurers to timely investigate and settle claims, to notify insured employees of the benefits and coverage contained in the insurance policies that are pertinent to a claim, and to refrain from attempts to obtain fraudulent releases of claims from their insureds?”¹⁸³ He notes several lower courts, which have held that state insurance bad faith remedies laws *do* regulate insurance and are thus saved from preemption.¹⁸⁴ Most of these cases, however, have subsequently been overruled since his article was published.

Bogan further argues that, taking Supreme Court precedent of saving insurance laws from preemption, those laws should be preempted because of the express language in section 514(b)(2)(B), even when the state remedies laws conflict with section 502, as the former expressly exempted state laws that regulate insurance.¹⁸⁵ The issue really boils down to two competing sections of the same federal statute: one provision expressly exempting certain state statutes from preemption, and one provision providing general remedies.¹⁸⁶ He asserts that statutory construction principles mandate that the courts give effect to legislative intent in such circumstances.¹⁸⁷

The savings clause is unambiguous in its exemption of state insurance laws. There is no support in the legislative history that the court may pick and choose which insurance laws to apply, especially with regards to state insurance remedies laws.¹⁸⁸ However, looking at the statement of Senator Williams on page 7 *infra*, it would appear that the remedies in section 502 eliminate all other laws, except for the provided “narrow exceptions” in ERISA. Bogan argues that the “narrow exceptions” *are* the state insurance bad faith laws, and thus are saved from preemption and section 502.¹⁸⁹ Therefore, as the Court has saved other insurance laws from preemption under the savings clause, it should also save state

¹⁸³ Bogan, *supra* note 6, at 134.

¹⁸⁴ *Id.* at 144-45. An example of the cases cited: *Colligan v. UNUM Life Ins. Co. of Am.*, No. 00-K-2512, 2001 U.S. Dist. LEXIS 8103 (D. Colo. Apr. 23, 2001), *overruled by* *Kidneigh v. UNUM Life Ins. Co.*, 345 F.2d 1182 (10th Cir. 2003).

¹⁸⁵ Bogan, *supra* note 6, at 152.

¹⁸⁶ *Id.* at 153.

¹⁸⁷ *Id.* at 153-54.

¹⁸⁸ *Id.* at 154 (citing *Metro. Life Ins. Co. v. Mass.*, 471 U.S. 724, 745-46 (1985)).

¹⁸⁹ *Id.* at 154.

insurance unfair practices statutes, as they were intended to be saved by Congress because they integrally affect the insurer-insured insurance policy relationship.

This is a different view on ERISA and does not entail rejecting any statutory language. I support this reading of ERISA; however, the Court has, as aptly described by Johnson, “painted itself into a corner” and likely will not adopt this reading. The Court is unlikely to undertake a sweeping change of heart and overrule twenty-plus years of precedent.¹⁹⁰

3. Alternative Pleading: One Argument to Use RICO for Fraudulent Claims Practices

In an article written by an employee benefits consultant, Lalena Turchi empathizes with the strife many beneficiaries face with benefits claims against fiduciaries.¹⁹¹ She is able to give a first-hand account of the administrative inefficiencies inherent in the insurer-provider billing system which leads to erroneous rejections, and the disheartening rejection due to a decision that the treatment is not “medically necessary.”¹⁹² She argues that, since the Court has rejected application of state insurance bad faith laws, the Racketeer Influenced and Corrupt Organizations Act (RICO), found in 18 U.S.C. § 1961, should be used.¹⁹³ It is of note that RICO claims allow for treble damages against defendants.¹⁹⁴

The Court allowed the plaintiffs-beneficiaries to assert a RICO claim in *Humana v. Forsyth*, and permitted the treble damage provision to

¹⁹⁰ Perhaps the Court only takes such drastic action once a millennium, such as was necessary in *Brown v. Bd. of Educ.*, 347 U.S. 483 (1954), and such action is not deemed as important for the health insurance industry and protection of plan participants.

¹⁹¹ Turchi, *supra* note 157, at 526.

¹⁹² *Id.* at 527.

¹⁹³ *Id.* at 551.

¹⁹⁴ See 18 U.S.C. § 1964(c): “Any person injured in his business or property by reason of a violation of section 1962 of this chapter may sue therefor in any appropriate United States district court and shall recover threefold the damages he sustains and the cost of the suit, including a reasonable attorney's fee, except that no person may rely upon any conduct that would have been actionable as fraud in the purchase or sale of securities to establish a violation of section 1962. The exception contained in the preceding sentence does not apply to an action against any person that is criminally convicted in connection with the fraud, in which case the statute of limitations shall start to run on the date on which the conviction becomes final.”

apply in a suit against an insurer involved in a scheme where it received discounts for hospital services but failed to pass those discounts on to the plan beneficiaries. The plan agreement, which provided that the insurer would pay 80% and the insured is responsible for the remaining 20%, was not followed, as the insurer paid at a discount and forced the insured to pay more than the agreed-upon 20%.¹⁹⁵ The Court held that RICO complemented the state statutory and common law claims for relief, and therefore did not supersede, preclude or conflict with the state laws under McCarran-Ferguson.¹⁹⁶ The claims alleged violations of RICO through a pattern of racketeering activity consisting of mail, wire, radio and television fraud.¹⁹⁷ The Court noted that RICO advances the state's interest in protecting against insurance fraud and does not frustrate state policy, and insurers have relied on RICO when bringing fraud claims.¹⁹⁸ However, this insurer was not governed under ERISA.

Turchi finds support for RICO claims in ERISA plans in several cases, such as *Maio v. Aetna*, where plan participants in an HMO brought suit against the insurer for overpayment of insurance.¹⁹⁹ While the court dismissed for lack of standing, it stated that the plaintiffs would have had standing if they had alleged that the health care received under the plan "actually was compromised or diminished as a result of insurer's management decisions challenged in the complaint."²⁰⁰ Accordingly, had the proper allegations been made, a RICO claim apparently would have been permissible. In another case, a plaintiff sued for a fraudulent insurance telemarketing scheme when an insurer sold her a death and dismemberment policy with limited value, and led her to believe it was a term life plan.²⁰¹ The district court, granting a motion to dismiss in part, however, stated that her loss of funds through premiums paid to the insurer was a cognizable injury to property as required under RICO.²⁰²

In *Klay v. Humana, Inc.*, a suit brought by physicians against HMOs, the doctors alleged that the HMOs were engaged in a scheme to underpay the physicians via a computer program that reimbursed the

¹⁹⁵ *Id.* at 552 (citing *Humana, Inc. v. Forsyth*, 525 U.S. 299, 299-301 (1999)).

¹⁹⁶ *See* 525 U.S. at 311.

¹⁹⁷ *Id.* at 299 (citing RICO, 18 U.S.C. § 1961).

¹⁹⁸ *Humana*, 525 U.S. at 299, 302.

¹⁹⁹ *Turchi*, *supra* note 157, at 554 (citing *Maio v. Aetna, Inc.*, 221 F.3d 472 (3d Cir. 2000)).

²⁰⁰ *Id.* (citing *Maio*, 221 F.3d at 472).

²⁰¹ *See McClain v. Coverdell & Co.*, 272 F. Supp. 2d 631, 634 (E.D. Mich. 2003).

²⁰² *Turchi*, *supra* note 157, at 555 (citing *McClain*, 272 F. Supp. 2d at 637).

physicians based on “financially expedient cost and actuarial data rather than medical necessity.”²⁰³ The RICO claims included racketeering activities through mail fraud, wire fraud, and extortion, as the HMOs threatened that the physicians, if they refused to cooperate, would lose patients, be blacklisted, and not be paid in full if they were not under contract with the HMO.²⁰⁴ The Eleventh Circuit held that “it would be unjust to allow corporation to engage in rampant and systemic wrongdoing, and then allow them to avoid class actions because the consequences of being held accountable for the misdeeds would be financially ruinous” in response to the insurers’ argument that such a suit would devastate the health care industry.²⁰⁵

Turchi argues that health plan participants suffer the same harm and experience the same calculated wrongdoing when denied benefits as the plaintiffs in the above mentioned cases, and RICO claims are the best method to obtain adequate remedies.²⁰⁶ Her analysis states that, in allowing a RICO claim, the Court will provide the much needed relief for ERISA preemption when an insurer denies benefits and insured suffers injuries, as it will force insurance companies to be more conservative in their benefit denials when there is a possibility for treble damages for bad faith.²⁰⁷

This approach is optimistic and inventive. Unfortunately, it is likely to be rejected by the Supreme Court, under the view that the remedies provided in section 502 are the sole remedies available. It is not an issue of federal versus state laws that can apply; the issue is that the language in section 502 is narrowly construed by the Court, and such a narrow interpretation will likely not allow a plaintiff to invoke RICO as the “other appropriate equitable relief” allowed under section 502(a)(3). An alternative to this is stricter regulatory oversight of insurers, specifically ERISA insurers, which should have the same result – compliance with insurance regulations and protection of beneficiaries.

IV. CASE STUDY: UNUM PROVIDENT

The Unum scandal has been immortalized as probably the most egregious known scheme to defraud participants and commit fiduciary

²⁰³ *Id.* at 557 (citing *Klay v. Humana, Inc.*, 382 F.3d 1241, 1246 (11th Cir. 2004), *cert. denied*, 543 U.S. 1081 (2005)).

²⁰⁴ *Id.*

²⁰⁵ *Id.* at 559 (citing *Klay*, 382 F.3d at 1274).

²⁰⁶ *Id.* at 560.

²⁰⁷ *Id.*

breaches. Unum is a disability insurer. It is made up of the parent company, Unum Group, and several subsidiaries, including Unum Life Insurance Company of America, the Paul Revere Life Insurance Company, Provident Life and Accident Insurance Company, Provident Life and Casualty Insurance Company, and First Unum Life Insurance Company. In 1997, the parent company acquired the Paul Revere Company, a Massachusetts corporation providing individual long-term disability insurance.²⁰⁸ In 1999, the parent company became Unum as a result of a merger between Unum Corporation and Provident Companies, Inc. When the companies were merged, an oversight resulted in no streamlining of policies, practices or procedures among the newly acquired groups.²⁰⁹ Thus, claims adjusters were localized and followed whatever procedures had been in place before the merger and had little oversight by the parent company.

A. THE LAWSUITS AND SCANDAL

An attorney at Unum realized that the company could gain discretionary review of its plan decisions, i.e., its claim denials, by incorporating *Firestone* language, which limits judicial review of benefit denials to only abuse of discretion situations.²¹⁰ An internal Unum executive memorandum advocated the “enormous advantages that ERISA, as interpreted by the courts, bestowed upon Unum” because of state law

²⁰⁸ MULTISTATE MARKET CONDUCT EXAMINATION REPORT, Rackeman, Sawyer & Brewster, P.C., submitted to state departments of insurance at 2 (Apr. 14, 2008).

²⁰⁹ Telephone interview with senior management employee, Unum (July 1, 2010). The interviewee requested to remain anonymous.

²¹⁰ *Id.* at 1321; see *Firestone Tire and Rubber Co. v. Bruch*, 489 U.S. 101, 111, 115 (1989) (Stating: “Trust principles make a deferential standard of review appropriate when a trustee exercises discretionary powers. See Restatement (Second) of Trusts § 187 (1959) ([w]here discretion is conferred upon the trustee with respect to the exercise of a power, its exercise is not subject to control by the court except to prevent an abuse by the trustee of his discretion’). See also G. Bogert & G. Bogert, *Law of Trusts and Trustees* § 560, pp. 193-208 (2d rev. ed. 1980). A trustee may be given power to construe disputed or doubtful terms, and in such circumstances the trustee’s interpretation will not be disturbed if reasonable. *Id.*, § 559, at 169-71. Whether ‘the exercise of a power is permissive or mandatory depends upon the terms of the trust.’ 3 W. Fratcher, *Scott on Trusts* § 187, p. 14 (4th ed. 1988).”)

preemption.²¹¹ Further, it advocated that such protection precludes jury trials, compensatory or punitive damages, and provides relief only for the amount of the benefit in question, which Unum would have had to pay anyway, but not before the claimant pays a lot of money and time to force such payment.²¹² The memorandum went on to state that twelve claims situations were identified “where we settled for \$7.8 million in the aggregate. If these 12 cases had been covered by ERISA, our liability would have been between zero and \$0.5 million.”²¹³ This memo illustrates ERISA’s weaknesses and insurers’ ability to exploit those weaknesses (loopholes which shield insurers from liability and penalties), to the detriment of beneficiaries with claims.

J. Harold Chandler, the CEO of Unum until 2003, instituted cost-containment measures, whereby claims processors were pressured to deny valid claims, especially during the last month of the quarter to meet “projections” and “budget goals.”²¹⁴ Several investigative reporting television programs demonstrated these mechanisms through internal Unum emails, which advised claims employees to deny claims in order to meet desired goals.²¹⁵ One Unum employee, a staff physician, admitted that Unum instructed him to use language to support denials of disability claims and denied him the ability to request additional medical testing to fully determine a claimant’s disability.²¹⁶

In *Weiss v. First Unum Life Insurance Co.*, the beneficiary, Weiss, sued First Unum under RICO, claiming that First Unum discontinued payment of his disability benefits as part of its racketeering scheme involving an intentional and illegal policy of rejecting expensive payouts to disabled insureds.²¹⁷ Weiss, an investment banker, had insurance through his employer, which included long-term disability (“LTD”) insurance by First Unum. He had a heart attack in 2001, which left him suffering permanently with ventricular tachycardia and unable to work due to lightheadedness, weakness, and shortness of breath.²¹⁸ His claim for was

²¹¹ *Id.* at 1321 (citing Memorandum from Jeff McCall to IDC Management Group & Glenn Felton, Provident Internal Memorandum, Re: ERISA (Oct. 2, 1995)).

²¹² *See id.*

²¹³ *Id.* (citing Memorandum, *supra* note 211).

²¹⁴ Langbein, *Trust Law as Regulatory Law*, *supra* note 6, at 1318.

²¹⁵ *Id.* at 1319.

²¹⁶ *Id.* (citing *McSharry v. Unum/Provident Corp.*, 237 F. Supp. 2d 875, 877 (E.D. Tenn. 2002)).

²¹⁷ *Weiss v. First Unum Life Ins. Co.*, 482 F.3d 254, 255-56 (3d Cir. 2007).

²¹⁸ *Id.* at 256.

approved for short-term disability benefits, which resulted in seven months of benefit payments.²¹⁹ At the end of the maximum allowable short-term disability benefits, Unum paid Weiss LTD benefits for another three months, and then discontinued paying benefits.²²⁰ Weiss brought suit in state court, and Unum removed to federal court and filed a motion to dismiss, arguing that the state claims were preempted under ERISA.²²¹ While the case was pending, Unum resumed payment of his LTD benefits, retroactive to the prior termination date.²²²

Weiss added both federal and state RICO claims, arguing that his claim was terminated because it exceeded \$11,000 per month.²²³ In support of this argument, he alleged that on October 3, 2001, defendants David Gilbert, Paul Keenan, George DiDonna, Lucy-Baird Stoddard, Unum employees, and others conspired at a roundtable meeting to terminate Weiss's benefits and devise a rationalization for doing so.²²⁴ Weiss claimed that DiDonna, the insurer physician, did not receive or examine his hospital records until the termination decision was reached, and tests that would make clear the severity of his injury were purposely never ordered.²²⁵ Weiss sidestepped a simple state bad faith claim by arguing that his denial is one instance in a pattern of fraudulent activity by First Unum aimed at depriving its insureds with large disability payouts of their contractual benefits.²²⁶

The District Court dismissed his RICO claims, believing that the allowance of such a RICO claim would interfere with New Jersey's statutory regulation of insurers, and thus run afoul of the McCarran-Ferguson Act.²²⁷ The Third Circuit Court of Appeals reversed, using the *Humana* analysis for ability to assert RICO claims.²²⁸ The court noted that

²¹⁹ *See id.*

²²⁰ *Id.*

²²¹ *Id.*

²²² *Id.*

²²³ *Weiss*, 482 F.3d at 257.

²²⁴ *Id.*

²²⁵ *Id.*

²²⁶ *Id.*

²²⁷ *Id.* at 255.

²²⁸ *Id.* at 256. The court expansively noted *Humana v. Forsyth* in its analysis, stating: "In sum, the *Humana* analysis explored the specific interplay between RICO and the state insurance scheme. As described above, the non-exclusive list of factors the Court examined in *Humana* included the following: (1) the availability of a private right of action under state statute; (2) the availability of a common law right of action; (3) the possibility that other state laws provided

the state insurance trade practices act (“ITPA”) allowed the insurance commissioner to determine whether an insurer had engaged in an unfair or deceptive practice, which includes unfair claims settlement practices.²²⁹ If a violation was found, the commissioner may assess a \$1,000 penalty for each negligent act or violation, unless the insurer knew or should have known the act was a violation - then the penalty is \$5,000 per act or violation.²³⁰ The commissioner was required to investigate upon receipt of a consumer complaint of a violation of this act, and upon a finding of violations, may order the insurer to pay restitution to the aggrieved party or other equitable relief.²³¹ The court also found there was a state common law private right of action against insurers for wrongly withheld benefits, other state laws allow claims, and the New Jersey Consumer Fraud Act, which provided treble damages available to redress such violations.²³²

The appeals court found that the claim alleged that First Unum embarked on a fraudulent scheme to deny insureds their rightful benefits, “clearly an unconscionable commercial practice in connection with the performance of its obligations subsequent to the sale of merchandise, i.e. payment of benefits.”²³³ Ultimately, the court held that Weiss could appropriately bring a RICO claim against First Unum, and remanded the case.²³⁴

In *McCauley v. First Unum*, the plaintiff’s disability insurance was through his employer plan and First Unum was both the administrator and payor of benefits.²³⁵ McCauley was diagnosed with advanced colon cancer in April 1991 and underwent severe chemotherapy treatments to save his life.²³⁶ Due to these treatments, he took several short-term disability leaves

grounds for suit; (4) the availability of punitive damages; (5) the fact that the damages available (in the case of Nevada, punitive damages) could exceed the amount recoverable under RICO, even taking into account RICO’s treble damages provision; (6) the absence of a position by the State as to any interest in any state policy or their administrative regime; and (7) the fact that insurers have relied on RICO to eradicate insurance fraud.” *Id.* at 261 (citing *Humana*, 525 U.S. at 311-14).

²²⁹ *Weiss*, 462 F.3d. at 263 (citing N.J. STAT. § 17:29B-5 (2003)).

²³⁰ *Id.* (citing N.J. STAT. § 17:29B-7(a) (2003)).

²³¹ *Id.* at 264 (citing N.J. STAT. § 17:29B-18 (2003)).

²³² *Id.* at 264-67.

²³³ *Id.* at 266.

²³⁴ *Id.* at 269. Note here that the court completely failed to mention ERISA’s preemption clauses and section 502, which limits the remedies available.

²³⁵ *McCauley v. First Unum Life Ins. Co.*, 551 F.3d 126, 129 (2d Cir. 2008).

²³⁶ *Id.* at 128-29.

in 1991 under the group disability plan.²³⁷ Over the next three years, he experienced health problems in connection with his cancer and took several short-term disability leaves.²³⁸ In 1994, he notified his employer that his health conditions were too traumatic and he could not continue to work.²³⁹

In May 1995, First Unum denied his claim.²⁴⁰ Upon McCauley's appeal and submission of additional information, First Unum again rejected his claim in September 1995.²⁴¹ McCauley attempted to rejoin the workforce and accepted a General Counsel position, but his health problems persisted.²⁴² Because he found work, his former employer ceased paying the disability plan premiums and advised him to convert the policy and make future payments, which he did.²⁴³ On January 16, 1996, he applied for long-term disability benefits under his conversion policy, as he accepted that he simply could not work due to his severe health problems.²⁴⁴ First Unum again denied his claim on the basis that the employment with the former employer had terminated in 1994, and, therefore, that he had exercised his conversion after the allowable period.²⁴⁵

McCauley brought suit for wrongful denial of benefits. He alleged that certain procedural irregularities, such as missing documents from his file and the incorrect assertion to McCauley that a medical doctor reviewed his file (only a nurse reviewed it) which occurred in the handling of his claim demonstrated that First Unum's conflict of interest had affected its decision to deny him benefits.²⁴⁶ While both courts held that the insurer's first denial of benefits was proper, as McCauley's physician's letter did not indicate total disability, the additional information submitted should have resulted in coverage.²⁴⁷ This additional information was a memorandum

²³⁷ *Id.* at 129.

²³⁸ *Id.*, noting, for instance, he had part of his liver removed because cancer was found, and subsequently suffered a severe liver infection. He also underwent surgery to repair a hernia. *Id.*

²³⁹ *Id.* The court noted his physical traumas, stating that, while the cancer treatment was successful, it caused "chronic diarrhea, chronic and acute renal impairment, incontinence, progressive vascular sclerosis, high cholesterol, insomnia, depression, and incisional scarring and pain." *Id.*

²⁴⁰ *Id.*

²⁴¹ *McCauley*, 551 F.3d at 129.

²⁴² *Id.*

²⁴³ *Id.*

²⁴⁴ *Id.*

²⁴⁵ *Id.*

²⁴⁶ *Id.* at 131.

²⁴⁷ *McCauley*, 551 F.3d at 134.

written by McCauley, with the advice and knowledge of his physician, listing his medical issues as (1) chronic diarrhea, (2) chronic and acute renal impairment, (3) progressive vascular sclerosis, (4) high cholesterol, (5) insomnia, and (6) incisional scarring and pain, and stated:

[He] is only able to control bowel movements by carefully timing his food ingestion and lists a number of ways in which this limits his daily activities. Respecting his renal impairment, the memorandum explains that McCauley has chronic blood in the urine and pain in the kidney area and that he forms a kidney stone every two weeks. As a result, his physician recommends that he not sit for long periods of time. Moreover, the memorandum states that during the acute phase of his renal impairment, “it is impossible for the patient to perform at any level.” As to his vascular sclerosis, the memorandum explains that McCauley’s vascular system was permanently damaged by the chemotherapy treatments and that he suffers “severe chronic headaches at the base of the skull, resulting in an inability to focus eyesight and a lack of concentration.” His insomnia is described as “chronic and recurring,” resulting in a “general feeling of lethargy and malaise” and leaving him with a “need to take naps during the day.” The memorandum also states that McCauley “is in pain on an almost constant basis” and takes Percocet, an opiate, to manage that pain.²⁴⁸

A nurse, not a doctor, reviewed this additional evidence and rejected his claim, because it was ““not an official document from [an] attending physician.””²⁴⁹ However, the rejection letter stated that it had rejected the claim because ““these conditions were acknowledged by your physician on the initial application and in his narrative letter of March 1995.””²⁵⁰

The court granted First Unum’s motion for summary judgment, stating that a *de novo* standard of review of the benefit denial was not applicable because the plan had *Firestone* language (which granted the insurer discretionary authority in benefit determinations, and a court may

²⁴⁸ *Id.* at 134-35.

²⁴⁹ *Id.* at 135.

²⁵⁰ *Id.* “First Unum never told McCauley that the absence of a physician’s signature was a reason for rejecting his information.” *Id.*

only review where there has been an arbitrary or capricious action).²⁵¹ On appeal, the Second Circuit Court of Appeals reversed, under the *Metropolitan Life v. Glenn* standard, and found that the plan administrator abused its discretion in denying the plaintiff's second claim for long-term disability benefits, because its reason for doing so was deceptive and unreasonable.²⁵²

Several bad faith suits were brought against Unum, and in one case (a non-ERISA case), a five million dollar punitive judgment was awarded by the jury due to the egregious bad faith acts by Unum.²⁵³ A district court in Massachusetts wrote that "an examination of cases involving First Unum . . . reveals a disturbing pattern of erroneous and arbitrary benefits denials, bad faith contract misinterpretations, and other unscrupulous tactics."²⁵⁴ That court listed more than thirty cases in which First Unum's denials were found to be unlawful, including one decision in which the behavior was "culpably abusive."²⁵⁵ A state insurance commissioner noted that Unum looked "for every technical legal way to avoid paying a claim."²⁵⁶ Several insurance state commissioners enforced fines, to the tune of millions of dollars, for the unfair and egregious claims practices of Unum employees.²⁵⁷

B. THE REGULATORY AUTHORITIES STEP IN

On January 7, 2003, the Massachusetts Division of Insurance conducted a market conduct examination of the Paul Revere Company's handling practices of individual disability insurance claims ("IDI"

²⁵¹ *Id.* at 130-31.

²⁵² *Id.* at 128, 133, 135 ("Following *Glenn*, a plan under which an administrator both evaluates and pays benefits claims creates the kind of conflict of interest that courts must take into account and weigh as a factor in determining whether there was an abuse of discretion, but does not make *de novo* review appropriate." (citing *Metro. Life Ins. Co. v. Glenn*, 128 S. Ct. 2348, 2348 (2008))).

²⁵³ Langbein, *Trust Law as Regulatory Law*, *supra* note 6, at 1319 (citing *Hangarter v. Paul Revere Life Ins. Co.*, 236 F. Supp. 2d 1069, 1082 (N.D. Cal. 2002), *aff'd*, 373 F.3d 998 (9th Cir. 2004)).

²⁵⁴ *Radford Trust v. First Unum Life Ins. Co of Am.*, 321 F. Supp. 2d 226, 247 (D. Mass. 2004), *rev'd on other grounds*, 491 F.3d 21, 25 (1st Cir. 2007).

²⁵⁵ *Id.* at 247 n.20.

²⁵⁶ Langbein, *Trust Law as Regulatory Law*, *supra* note 6, at 1320 (citing Mike Pare, *\$1 Million Fine Hits Unum*, CHATTANOOGA TIMES FREE PRESS, Mar. 19, 2003, at C1).

²⁵⁷ *See id.*

claims).²⁵⁸ In September 2003, a multistate targeted market conduct examination was commenced by the Maine Bureau of Insurance, the Massachusetts Division of Insurance, and the Tennessee Department of Commerce and Insurance, concerning the claims practices of Unum, Revere and Provident in both IDI and group LTC policies.²⁵⁹ Additionally, the remaining forty-seven states plus the District of Colombia acted as “participating states” in this 2003 exam.²⁶⁰ Contemporaneously with the Multistate Examination, the Department of Labor conducted an investigation of the companies (the “DOL Investigation”) pursuant to Section 504 of ERISA.²⁶¹

According to the exam report, the examination team requested the companies to provide a comprehensive database including all claims closed during 2002.²⁶² Initially, 300 claim files randomly selected from IDI and LTD claims closed during 2002, or for which benefit determinations were appealed or litigated during 2002, or claims open as of year-end 2002 were evaluated for the initial review.²⁶³ The initial review comprised 300 claims (100 claims each for Unum, Revere and Provident).²⁶⁴ “The proportion of selected IDI and LTD claims was based on the relative reported reserves for each company as of December 31, 2002.”²⁶⁵ The team also commenced a second, follow-up review in 2004, as the companies advised that they had made several changes in claim administration implemented in 2003.²⁶⁶

The initial review actually consisted of 299 files instead of 300, because the companies were unable to locate one of the claims files

²⁵⁸ 2008 MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 208, at 3.

²⁵⁹ *Id.*

²⁶⁰ *Id.* at 3-4.

²⁶¹ *Unum Provident Regulatory Settlement Agreement*, § A.3, MAINE.GOV (Nov. 2004), http://www.maine.gov/pfr/insurance/unum/UNUM_Regulatory_Settlement_Agreement.htm.

²⁶² TARGETED MULTISTATE MARKET CONDUCT EXAMINATION REPORT, MAINE.GOV (Nov. 18, 2004), http://www.maine.gov/pfr/insurance/unum/Unum_Multistate_ExamReport.htm.

²⁶³ *Id.*

²⁶⁴ *Id.*

²⁶⁵ *Id.*

²⁶⁶ *Id.* In 2004, the team reviewed 75 claim files (25 each for Unum, Revere and Provident) which were randomly selected from the Companies’ IDI and LTD claims for which benefit determinations were first appealed during the period of December 2003 through February 2004. *Id.*

selected for review.²⁶⁷ This initial review concluded with four “general areas of concern” for both the IDI and LTD claims handling. The four areas were: (1) excessive reliance on in-house medical professionals; (2) an unfair bias and inappropriate interpretation of medical reports to the detriment of claimants based on the excessive reliance upon the in-house medical professionals; (3) failure to evaluate the totality of the claimants’ medical conditions (benefits were denied due to the failure to “properly evaluate cumulative effects” of multiple claimant conditions); and (4) the inappropriate burden placed on claimants to justify eligibility for benefits.²⁶⁸

After the follow up review in 2004, the team concluded that further regulatory action was necessary, resulting in the Regulatory Settlement Agreements (“RSA”), which provided a “Plan of Corrective Action” that the companies implement in their claims handling procedures.²⁶⁹ The RSA stated that a further review would be completed in 24 months to assess implementation of the practice and procedures set forth in the RSA.²⁷⁰ The RSA also provided for a \$15 million penalty to be paid by the companies.²⁷¹

The RSA required the following changes in claims practices and procedures to reduce any potential bias and promote claims handling accuracy:

- The engagement of experienced claim personnel at the earliest possible stage of claim reviews;
- Increased emphasis upon claim staff accountability for compliance with the terms of insurance policies and applicable law;
- Increased involvement of higher levels of claim management staff in each claim denial or claim termination decision;
- Creation of a separate compliance/accountability function at the claim denial and claim termination level;

²⁶⁷ *Id.*

²⁶⁸ TARGETED MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 262.

²⁶⁹ *Id.*

²⁷⁰ *Id.*

²⁷¹ *Id.*

- Assurance that co-morbid conditions are properly evaluated at every level of claim review;
- Increased utilization of Independent Medical Examinations;
- Additional compliance training for all claim staff, with emphasis upon the results of the multistate examination, the Plan, and the NAIC Unfair Claim Settlement Practices Act; and
- Additional training for group policyholder human resources personnel so as to better facilitate the process for LTD claims.²⁷²

With regards to the corporate governance, the companies were required to address corporate control issues by implementing the following changes:

- The Board of Directors of the Parent Company will be expanded by three members, each of which will have significant insurance industry or insurance regulatory experience (two will have regulatory experience); each candidate will be approved by the Lead States and by the DOL;
- The Audit Committee of the Board of Directors will be expanded by one member; at least one of the new members of the Board of Directors will be appointed to the Audit Committee;
- The Board of Directors will establish a new Regulatory Compliance Committee, comprised of two of the new members of the Board, and three existing independent directors; the Regulatory Compliance Committee will have responsibility for monitoring compliance with the Plan and other compliance-related oversight functions; and
- The companies will create a Regulatory Compliance Unit, which will report directly to the Regulatory Compliance Committee; the Regulatory Compliance Unit will monitor compliance with the Plan (including the functions of the Claim Reassessment Unit) through the performance of periodic audits, provide assistance

²⁷² *Id.* See also Unum Provident Regulatory Settlement Agreement, *supra* note 261, at § B.3.h.

to claimants to ease and facilitate the claim submission process, and gather data for the Lead States' ongoing monitoring of compliance with the Plan.²⁷³

Further, the RSA and Plan required the Regulatory Compliance Committee, the companies' senior management, the lead regulators and the DOL to meet on a quarterly basis to evaluate compliance with the Plan and RSA.²⁷⁴ In 2005, the RSA was amended to allow LTD and IDI claimants the opportunity to have denied claims reassessed.²⁷⁵

The reassessment of claims was performed by the Claim Reassessment Unit ("CRU"), a newly formed claims unit within the companies.²⁷⁶ Its results were evaluated in another multistate market conduct examination, conducted in 2007, to evaluate compliance by the companies. More than 79,000 claimants elected to have their claims reassessed, however, only 23,190 claimants correctly submitted the required Reassessment Information Forms.²⁷⁷ Therefore, only 23,190 claims were actually reassessed by the companies, which is approximately 29% of the total requested reassessments.²⁷⁸

The examining team concluded that 41.7% of the claims (including both IDI and LTD) reassessed were reversed in whole or part, resulting in approximately \$676.2 million in benefits paid to claimants, either immediately or reserved for future payments.²⁷⁹ Forty-five percent of the LTD reassessed claims were reversed in whole or part, resulting in approximately \$558.6 million of benefits paid or reserved for future payment.²⁸⁰ Twenty-two percent of IDI reassessed claims were reversed in whole or part, resulting in approximately \$117.6 million in benefits paid or

²⁷³ TARGETED MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 262.

²⁷⁴ *Id.*

²⁷⁵ *Amendment to Unum Provident Regulatory Settlement Agreement*, MAINE.GOV (Oct. 3, 2005), http://www.maine.gov/pfr/insurance/unum/UNUM_RSA_Amendment.htm.

²⁷⁶ 2008 MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 208, at 6.

²⁷⁷ *Id.*

²⁷⁸ *Id.* It is unclear why only 29% of the claimants correctly filed the Reassessment Forms. The Report indicates that Unum correctly mailed the forms to the claimants, as required under the RSA.

²⁷⁹ *Id.*

²⁸⁰ *Id.*

reserved for future payment.²⁸¹ In total, it appears the companies paid more than a billion dollars in benefits due to the reassessed claims.

The 2007 exam also evaluated the companies' compliance with the claims procedures required under the RSA. The team evaluated claims reassessed by the CRU, both IDI and LTD, along with a sample of claims assessed by other claims personnel.²⁸² The team assessed 50 CRU IDI claims, 100 CRU LTD claims, 50 Operations IDI claims, and 100 Operations LTD claims.²⁸³ According to the RSA, the error rate in claims could not exceed 7% for each area assessed.²⁸⁴ The team concluded that the companies were well below this error rate, and in some instances there were zero errors.²⁸⁵ Therefore, according to the 2008 multistate market conduct examination report, the companies are in complete compliance with the RSA.

In 2007, California's Department of Insurance ("CDI") also conducted an independent market conduct exam, releasing the report in 2008.²⁸⁶ The examiners evaluated the reassessed claims subject to the separate California Settlement Agreement between Unum and CDI. One hundred and ninety-one reassessed claims were reviewed, along with 30 post-CSA claims (closed between December 2005 and May 2006) and 60 post-CSA claims (closed between August 2006 and July 2007).²⁸⁷

The examiners found seven violations in evaluating the 191 reassessed claims (approximately 3% error).²⁸⁸ The violations included: (1) five instances of failure to comply with the CSA's definition of "total disability" in denying claims and (2) two instances of failure to effectuate

²⁸¹ *Id.* at 7.

²⁸² 2008 MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 208, at 8-9.

²⁸³ *Id.* at 9.

²⁸⁴ *Id.* at 10; *see also* Unum Provident Regulatory Settlement Agreement, *supra* note 261, at § B.3.j.

²⁸⁵ 2008 MULTISTATE MARKET CONDUCT EXAMINATION REPORT, *supra* note 208, at 10.

²⁸⁶ CALIFORNIA DEPARTMENT OF INSURANCE, REPORT OF THE MARKET CONDUCT EXAMINATION OF THE CLAIMS PRACTICES OF UNUM LIFE INSURANCE COMPANY, PROVIDENT LIFE AND ACCIDENT INSURANCE COMPANY, AND PAUL REVERE LIFE INSURANCE COMPANY 2 (2008), *available at* http://www20.insurance.ca.gov/e_pubacc/REPORT/109890.htm [hereinafter CALIFORNIA INSURANCE MARKET CONDUCT REPORT].

²⁸⁷ *Id.* at 3. The breakdown of claims reviewed: 137 group LTD claims and 54 IDI claims. *Id.* at 7.

²⁸⁸ *Id.* at 4.

prompt, fair and equitable settlements of claims where liability was reasonably clear.²⁸⁹ In response, Unum voluntarily imposed a written refresher training course for the CRU employees, emphasizing compliance, and reassessed the noted claims.²⁹⁰

There were no violations found in the 90 post-CSA claims.²⁹¹ The report noted a 54% drop in consumer complaints against Unum after it adopted the measures required in the regulatory settlement agreements.²⁹²

C. THE REST OF THE STORY: UNUM IS NOW A MODEL DISABILITY INSURER AND TRIES TO AMELIORATE ITS BAD REPUTATION

I had the opportunity to speak with a senior management employee at Unum. He acknowledged the egregious behavior which occurred and resulted in the lawsuits and RSA, but was very clear about Unum's complete turnaround. He advised that it is a completely different company under these new claims practices and procedures, as evidenced by the most recent market conduct reports. The procedures implemented to ensure compliance and fairness include an ethics hotline, where employees are encouraged to report any wrongdoing in business practice and remain anonymous, a notice to beneficiaries upon a denied claim of their right to request an independent medical exam, a requirements for claims personnel to contact attending physicians if there are questions or when clarification is necessary, and a policy to give significant weight to social security disability decisions and attending physician decisions.²⁹³

The departments of insurance view this outcome as a success. In a 2009 letter from the Maine, Massachusetts, Tennessee and New York departments of Insurance to the editor of the Insurance Forum, the commissioners and superintendents advocated that the "systemic misconduct" that led to the multistate examinations was no longer present at Unum, according to the latest exam reports.²⁹⁴ Mila Kaufman, superintendent of the Maine Department Insurance declared that "this case . . . is an example of effective state-based insurance regulation for insurance

²⁸⁹ *Id.* at 10-11.

²⁹⁰ *Id.* at 6.

²⁹¹ *Id.* at 4.

²⁹² CALIFORNIA INSURANCE MARKET CONDUCT REPORT, *supra* note 286, at 7.

²⁹³ *Id.* at 6.

²⁹⁴ Letter from Me. Bureau of Ins., Mass. Div. of Ins., Tenn. Dep't of Commerce and Ins., and N.Y. Dep't of Ins. to Joseph Belth, Editor, The Ins. Forum (Jan. 6, 2009) (on file with author), *available at* http://www.maine.gov/pfr/insurance/unum/unum_exam_settlement.htm.

consumers. Regulators identified a problem and worked together to effectively address it.”²⁹⁵ She went on to support Unum’s change, stating “it is also an example of an insurer reforming its practices and becoming a model for other insurers. The strong new processes and the resulting change in corporate culture – measure by the very low rate and in some cases a 0% error in claim determinations is remarkable.”²⁹⁶ This view was also advocated by the Massachusetts Insurance Commissioner, stating she “is pleased” with Unum Group’s compliance and using the procedures to “ensure its claimants are treated fairly going forward.”²⁹⁷

In my interview, the company seems to be frustrated with the lack of knowledge of the RSA and changes in procedure. In my insurance classes, when studying bad faith and ERISA law, we read about the Unum cases but never read about the regulatory involvement; perhaps due to time constraints. Unum is a great case study to show regulators working together to reform an insurer into a fully compliant and better market actor. During the RSA negotiations, Unum was advised that the standards provided in the RSA would eventually be enforced nationally. Unfortunately, this is not true, as is seen from the Good Morning America cases.²⁹⁸

V. MORE REGULATORY OVERSIGHT WILL ALLEVIATE EGREGIOUS INSURER ACTIONS AND ENSURE COMPLIANCE

The kind of regulatory cooperation between states that occurred in the Unum case is exactly what is needed at present to ensure beneficiaries are protected from unscrupulous insurers. As discussed, ERISA fails to provide the remedies to claimants when benefits are wrongly denied. As the Supreme Court is unlikely to change its position on interpreting ERISA and Congress is unlikely to amend ERISA, it is left to the state regulators to

²⁹⁵ Press Release, Mila Kaufman, Superintendent, Me. Bureau of Ins., Unum Landmark Multi-State Settlement Agreement and Examinations Concluded (Apr. 15, 2008) (on file with author), *available at* <http://www.maine.gov/tools/whatsnew/index.php?topic=INS-PressReleases&id=54160&v=Default>.

²⁹⁶ *Id.*

²⁹⁷ Press Release, Mass. Office of Consumer Affairs & Bus. Regulation, Div. of Ins. Completes Review of Unum Group Disability Claims Payments (Apr. 15, 2008) (on file with author), *available at* http://www.mass.gov/?pageID=ocaterminal&L=6&L0=Home&L1=Government&L2=Our+Agencies+and+Divisions&L3=Division+of+Insurance&L4=Archive+of+DOI+News+%26+Updates&L5=2008+DOI+Press+Releases&sid=Eoca&b=terminalcontent&f=doi_Media_media_press_73&csid=Eoca.

²⁹⁸ *See supra* Part III.B, pp. 29-30.

effect change and compliance. Per the savings clause, ERISA insurers are still subject to compliance with state insurance laws. Thus, regulators must strictly enforce the types of procedures required of Unum as to all ERISA insurers, which will alleviate the need for money damages; as such procedures and examinations will ensure compliance and result in protection of beneficiaries. To this end, the National Association of Insurance Commissioners (NAIC) should assist in implementing this change.

A. PURPOSE OF STATE DEPARTMENTS OF INSURANCE

Each state has a regulatory authority which oversees the insurance industry transacting business in its state. They generally regulate insurer activity and compliance with all state insurance laws and regulations, such as licensing, policy form approval, rate approval, ensure adequate financial conditions, receive consumer complaints and conduct market conduct examinations to ensure compliance and fairness to consumers.²⁹⁹ These departments exist primarily to ensure compliance with state laws and to protect consumers. ERISA-insured plans are governed by these laws and oversight by the state departments of insurance.

B. REGULATORY COOPERATION WILL LEAD TO INCREASED BENEFICIARY PROTECTION

As demonstrated above, the Supreme Court has held that ERISA does not allow compensatory or consequential money damages beyond that of the denied benefit. In *Mass. Mutual v. Russell*, the U.S. Supreme Court held that an unreasonable delay in receipt of benefits does not warrant consequential damages, or money damages, being paid to the beneficiary in an individual capacity, as ERISA only envisioned the plan to receive such damages.³⁰⁰ In *Varity Corp.*, the Court held that an individual may recover

²⁹⁹ See, e.g., Conn. Dep't of Ins., *Purpose and Duties of the Connecticut Insurance Department*, CT.GOV, <http://www.ct.gov/cid/cwp/view.asp?q=254384> (last visited July 4, 2010); *State Insurance Regulation: History, Purpose and Structure*, NAIC, available at http://www.naic.org/documents/consumer_state_reg_brief.pdf (last visited July 4, 2010); Ind. Dep't of Ins., *Indiana Department of Insurance*, IN.GOV, <http://www.in.gov/idoi/> (last visited July 4, 2010).

³⁰⁰ *Russell*, 473 U.S. at 144, 148.

individually.³⁰¹ In *Pilot Life* and *Andrews-Clarke*, the Court refused to apply state bad faith statutes because they were not specific to the insurance industry, no matter how egregious the misconduct by the insurer.³⁰² In *Aetna Health, Inc. v. Davila*, the Court refused to apply a state insurance bad faith statute to the HMO, noting that although it was specific to the insurance industry, an HMO's mixed decision on treatment and eligibility was not a fiduciary function, thus it could not be held liable for any damage that occurred as a result.³⁰³ The Court instead suggested that beneficiaries who were wrongly denied benefits pay for the services out of their own pocket and then bring suit for enforcement.³⁰⁴ This statement by the Court is completely out of touch with America and current economic conditions. With unemployment at 10% and economic crisis abroad, what average person, who relies on an employer-provided health plan, has the wherewithal to single-handedly pay for medical services, which are extremely expensive, out of their own pocket? As shown in the Good Morning America cases, several of the beneficiaries who were denied benefits were left with no income and often choose between paying the mortgage, the utilities, or food for their family, never mind paying an attorney to fight with the insurer over the denied benefits.³⁰⁵

As a result, there is only one authority left with the ability to adequately regulate this industry and protect consumers: the state departments of insurance. Not only is it their stated purpose as state agencies, but it is sorely needed due to the lack of oversight elsewhere. The Unum example demonstrates that regulators are able to cooperate and work together to implement fair practices and procedures and oversee compliance. The state departments of insurance simply need to enact the same practices and procedures from the RSA in each state, and even on a national level, to ensure beneficiaries everywhere are protected from insurer opportunism.

³⁰¹ See 29 U.S.C. § 1132(a) (2011); *Varity Corp. v. Howe*, 516 U.S. 489, 510 (1996).

³⁰² *Pilot Life Ins. Co. v. Dedeaux*, 481 U.S. 41, 50 (1987); *Andrews-Clarke v. Travelers Ins. Co.*, 984 F. Supp. 49, 55-56 (D. Mass. 1997).

³⁰³ *Aetna Health Inc. v. Davila*, 542 U.S. 200, 212 (2004).

³⁰⁴ *Id.* at 211.

³⁰⁵ *Cuomo & Wagschal*, *supra* note 145.

VI. CONCLUSION

The current regime of insurer regulation in the ERISA context is in danger of harming beneficiaries due to the lack of motivation by insurers to be fair and to strictly comply with state claims practices laws. To combat this inequity, I propose that the state departments of insurance intercede and enforce strict regulatory oversight of such insurers, as was done in the Unum situation. There, the state departments of insurance collaborated and implemented practices and procedures for Unum to adopt and use in its claims handling. As a result, Unum has drastically changed in its claims practices and is arguable a model for the industry in terms of its customer service of claimants. These practices and procedures are not nationally enforced. If such practices and procedures were enforced, it would admittedly serve as an added expense on insurers. However, these practices and procedures would more effectively regulate this area, thereby leading to fairness in claims handling, and ultimately protecting beneficiaries the entire goal behind ERISA. If strict regulation is in force, then beneficiaries would not need to resort to additional remedies, as the insurers would fear large penalties from state department of insurance for noncompliance. This fear, along with reputation damage, would keep ERISA insurers in line.

KEYNOTE ADDRESS, ACTUARIAL LITIGATION: HOW STATISTICS CAN HELP RESOLVE BIG CASES⁺

KENNETH R. FEINBERG^{*}

Thank you very much. The Dean said I'm a skilled mediator. As long as he didn't say I'm a model mediator. A couple of months ago, somebody introduced me as a model mediator, and a critic stood up and said, "Model in Webster's Dictionary; a small replica of the real thing." I want to thank the Dean for mentioning my book, *What is Life Worth*. Now you may have trouble finding that book today in Barnes & Noble or on Amazon.com. Don't worry, my personal supply of my book is virtually inexhaustible.

It's a real pleasure for me to be here today. There was no way I was turning down an opportunity to be here today and partake in a small part of [this symposium]. It is true, as the Dean pointed out, that for various reasons, which you can read in my book, I got involved in some of these public law challenges over the last twenty-five years starting with Judge Weinstein, Agent Orange, 9/11, Virginia Tech, the Pay Czar (which I think is a very unfortunate characterization of my role, Pay Czar).

[L]et me make a few comments about all of these assignments. First, these special funds that are set up, like the one I'm doing now, with BP, are very, very rare and they should be very rare. I'm asked all the time [whether these funds], like the 9/11 fund, the Agent Orange fund, the Virginia Tech fund, [are] "an alternative to the conventional way of resolving disputes, the wave of the future?" Absolutely not. They are not the wave of the future, they should not be the wave of the future, and even if you wanted them to be replicated, it won't happen. I mean I do these every six years when some tragedy befalls the country. There are variations; you heard . . . others talking about bankruptcy and mass settlements. Those are cousins to what I do. I'm engaged with a design and

⁺ *Connecticut Insurance Law Journal* Symposium, Actuarial Litigation: How Statistics Can Help Resolve Big Cases, April 15, 2011, Hartford, CT.

^{*} Kenneth R. Feinberg, Founder and Managing Partner, Feinberg Rozen LLP, Author of *WHAT IS LIFE WORTH?* (Public Affairs Press, 2005). Among his many accomplishments, Kenneth Feinberg is the Administrator of the BP Claims Fund and was the Special Master of the Federal September 11th Victim Compensation Fund.

an implementation and an administration that's very, very different from those in terms of its lack of checks and balances, in terms of its delegation, and in terms of its function.

What I do is really out of the box. It is a precedent to nothing and this BP thing is a wonderful idea of why it's a precedent for nothing. Within thirty days after the deep-water explosion, the administration sits down with BP; no statute, no judicial oversight, no regulation, [no] administration, no checks and balances in the form of senate confirmation or me or anybody like me, and they shake hands and BP announces to the world, "We have decided to put up twenty billion dollars." Now how many times can you think of a company, or an admitted wrongdoer, who decides that it's in our interest and it's in the public interest [to] fund twenty billion dollars?

[After] 9/11, Congress passed a law to set up a special fund. Congress didn't appropriate one dollar for that fund. Congress just said, "Whatever it costs, Feinberg, pay it out of petty cash from the US Treasury." No appropriation, pay whatever it takes to get the job done. [With the] Agent Orange [fund], under the auspices of Judge Weinstein, . . . eight chemical companies . . . decided to settle a class action . . . put up one hundred eighty million dollars with interest, which over ten years will grow to about two hundred fifty million dollars. That's when interest mattered . . .

These programs – ask yourself a very practical question when you think about these special programs. Who is funding these programs? Who is going to fund them? If you don't have a deep pocket willing to cut the check; even if you had a deep pocket, as a philosophic matter, these programs are so alien to the conventional way we resolve disputes that even if you come up with some sort of law review article that's going to point out how these programs can be expanded and made more pervasive along the legal spectrum; I don't think you'll get much support for it.

The conventional adversarial litigation system is so engrained in the fabric of this country. It's such a part of our history, our heritage; you may nibble at the edges, but you're not going to change that system. . . . There is a problem with any profession. There's a problem with the adversary system, but when you talk about wholesale changes, here's an alternative way to do it. Unless that alternative way is grounded in history as an alternative – and there are some examples, workers' compensation, the most obvious – unless you're going to have a real historical basis for making radical change, these programs that I've administered over the past twenty-five years should be viewed as aberrations, as one-offs.

Now people . . . may look at the way these programs are administered and designed and you may say, “You know, we can take a little of this and get a little of that.” That’s fine, but any idea that, “Well we did it for BP, we did it for 9/11, let’s do it for asbestos, or let’s do it for pharmaceuticals or let’s do it for chemicals,” it isn’t going to happen. That’s why you really need . . . these people . . . within the design mold of mass aggregation and mass ligation, class actions; a discussion of that, that’s where I think you’re going to likely see some change. That’s where you may see some change, not what I do, which is really an aberration and a sideshow.

Philosophically, you see, there are real problems with what I do. Now take 9/11; the 9/11 Victim Compensation. . . . It was not a mistake. The 9/11 Victim Compensation Fund was absolutely the right thing to do and I’ve been defending it since it was passed by Congress, signed by the President and we implemented it. It was successful and it was the right thing to do, but they’ll never do it again. They will never do it again, not the way it was done in 2001 by Congress where it delegated to one person the authority to design and administer a program that by statute gave everybody a different amount of money in order to buy them out – to attract them out of a tort system. And even though I’ll defend the 9/11 Program, it is a very, very close question. Philosophically, it raises tremendous issues in a free society. There was no 9/11 Fund for the victims of Katrina. You should have read some of the emails that I got when I was administering that fund.

“Dear Mr. Feinberg, my son died in Oklahoma City, where’s my check?” “Dear Mr. Feinberg, I don’t get it. My daughter died in the basement of the World Trade Center in the original 1993 attacks committed by the very same people, how come I’m not eligible by a check?” And it didn’t end with terrorism. “Dear Mr. Feinberg, explain something to me. Last year my wife saved three little girls from drowning in the Mississippi River and then she drowned a heroine. Where’s my check?”

It’s very, very difficult to justify public money, yet tax payer money one hundred percent to some people who are victims of life’s misfortune. Nobody else is getting a check, and even though I think it could be justified with 9/11 and it was the right thing to do, it’s not by trying to explain away differences among victims – I can’t do that. Maybe you can, but from the perspective of the country, it was the right thing to do, not from the perspective of the victims. The country wanted to do it. The country wanted to show its community and cohesiveness with the victims. Fine, from the country’s perspective, but very, very difficult issues get raised. Then to delegate to one person – I mean I’m getting hammered

now with BP and with the anniversary coming up next week – hammered. One person is making these decisions.

Now there are differences obviously with all of these programs. The problem in BP is the sheer volume paid. I’ve received, in nine months, eight hundred thousand claims. You know, its human nature. BP announces to the world, “Twenty billion. We’re going to make you whole.” Well, I’m a dentist, are they going to make me whole? I’m a veterinarian, a chiropractor, why shouldn’t I file a claim? Proximate cause, what is that? That’s something taught maybe at Connecticut Law School, but what that have to do with me? But for causation is what I’m interested in. But for the skill and human nature being what it is, every financial ill-eye suffers because of that bill. It’s in good faith and you’re not going to convince people otherwise. So in nine months, we have distributed to two hundred thousand people, four billion dollars. In nine months, and you still get criticized.

So these programs raise important issues, philosophical, political, etcetera. . . . How do you decide what a person ought to get paid, whether it’s in tort or [when] Congress passes a law that says, “Pay Czar.” How do you decide what a corporate official at CitiGroup should get paid? How do you decide compensation? How do you decide it?

Now the first thing you find out is when you read the statutes or the rules or the compact or the escrow agreement; that gives you some guidance as to how you’re going to go about compensating. Notice that in these cases, they largely involve death, physical injury. Pay Czar is financial compensation outside the tort system and BP is largely – not exclusively but largely – old fashioned economic loss: financial wage loss, income loss arising out of the spill. [There’s] very, very little actuarial work [involved], very little. First of all, you basically have an unlimited budget in a lot of these, you see, so there’s no actuarial requirement in terms of trying to allocate limited resources, thank goodness. One of the biggest problems that I confront in my work is when Peter thinks I’m nickel and diming him in order to pay Paul out of a limited fund, you see. So the more money I have, the more I’m able to deflect arguments like that.

Most of the work in defining compensation depends on the statute and the nature of the cohort of people you’re trying to help. In 9/11, the overwhelming compensation went to death. We paid about three thousand people something like six billion dollars. . . . Notwithstanding the title of my book, *What is Life Worth*, I like the title because it sells books, but it is a little bit of a misnomer . . . because we’re not placing value in 9/11 on the moral integrity of any person. I’m doing what judges and juries do in Connecticut every day. . . . What is the economic loss suffered by the

victim? How old was he? How long would he have worked? What does the census and the Bureau of Labor Statistics tell us about what a secretary makes in the World Trade Center? Male or female? How many dependents?

Put aside the administrative law problem of having one person do it and delegate to one person, it is basically rather common, if you want to know the truth. What I do in all of these assignments is not rocket science. The people in this room could do exactly what I do, exactly. This is not something where I have a magic bullet here. I do what judges and juries do every day. In BP, but for the difficult problem, shall we say, of people not paying their taxes so that it's hard to prove the claim, it's rather straightforward. What did you earn before the spill? What did you earn after the spill? How do you tie the difference to the spill?

I mean I've got claims from fifty states, from fifty states. "Mr. Feinberg, we served the best shrimp scampi in Hartford. Now we can't get shrimp from the Gulf and we've lost ten percent of our clientele base." "Well proximate cause. . ." "What? What does that mean? Pay me. You've got twenty billion dollars."

One final point . . . One thing that I'm not involved in, in any of these cases, is insurance, contribution, subrogation – not on my watch. Thank goodness. Here's a claimant. Pay the claimant or don't pay the claimant. "Well, there are offsets from insurance." . . . "There's an indemnity agreement that BP has with Transocean and with Halliburton." I don't want to hear about it. That's not what I'm not here for. I'm here to pay the claimant and get a release. Corral the claims and pay the people. Now if BP and Transocean have contributions, indemnity agreements, offsets, insurance, it's not my problem. My problem is to pay the claim. Whatever else is going to happen is going to happen, but not on my watch. Otherwise, I'm already bogged down with eight hundred thousand claims. People want their money, you see. . . . there have been two hundred seventy thousand claims filed since November 23rd, and about seventy percent of them have been processed. Not all paid, but processed.

So that's sort of what I'm doing, and why I thought . . . this [was] the right crowd to [consider] the ramifications of what I do or what our cousins do in the mass tort system and . . . aggregative statistics and aggregative law.

A NORMATIVE EVALUATION OF ACTUARIAL LITIGATION

ROBERT G. BONE*

This Article addresses the normative issues raised by the use of statistical sampling to adjudicate large case aggregations. In its recent decision, Wal-Mart Stores, Inc. v. Dukes, the Supreme Court referred to sampling pejoratively as “Trial by Formula.” This Article argues that the pejorative label is undeserved. In fact, sampling can be justified in many more situations than courts currently apply it, and society is paying a very high price for limiting its use. I explored some of the normative issues in an earlier publication, Statistical Adjudication: Rights, Justice, and Utility in a World of Process Scarcity, and the current Article expands on my earlier analysis in four respects. First, it analyzes the effect of sampling on settlement and discusses in more detail the problem of frivolous and weak filings. Sampling tends to reduce the likelihood of settlement and also provides cover for undesirable lawsuits. However, while both of these effects must be considered in any efficiency analysis, neither is likely to tip the cost-benefit balance against the use of sampling in large enough case aggregations. Second, this Article evaluates sampling in the context of an outcome-oriented rights-based theory. In this connection, the most serious problem is that sampling gives high value plaintiffs only an average recovery. Statistical Adjudication discussed this topic as well, but the current Article generalizes the analysis in a useful way. Third, the Article offers some further thoughts about process-based participation and the day-in-court right based on work that post-dates Statistical Adjudication. Fourth, the Article explores another possible objection to sampling that Statistical Adjudication did not address. This objection, which I call the “methodological legitimacy objection,” is distinct from adverse effects on outcome and limitations on individual participation. It rests ultimately on the assumption that adjudication at its core involves reasoned deliberation that engages the facts of particular cases. The problem with sampling from this perspective is that it substitutes a formulaic method for fact-sensitive reasoning. This Article shows that while the methodological legitimacy objection has some intuitive appeal, it is very difficult to sustain in a rigorous way.

I. INTRODUCTION

Statistical methods pervade the law and litigation. Sometimes the substantive law adopts an explicitly probabilistic standard suitable for evaluation by statistical techniques. One example is the likelihood of confusion requirement for trademark infringement, which calls on the court to estimate the probability that an ordinary consumer will be confused.¹ Sometimes the substantive law adopts a standard that, while not explicitly probabilistic on its face, nevertheless authorizes statistical methods. An example is the discrimination element of a disparate impact Title VII claim, which sometimes involves a statistical analysis to determine the existence and magnitude of differential effects.² Indeed, constitutional civil rights claims based on the Equal Protection Clause often depend on statistical evidence insofar as liability turns on a comparison of the challenged conduct to statistical features of a larger population.

Statistical methods are used even more frequently to generate evidence to prove a claim. Sometimes a legal standard that is not inherently linked to statistical properties of a phenomenon is nevertheless most easily proved statistically. For example, liability might turn on acts of the defendant reflected in written records too numerous to examine individually. In such a case, the plaintiff might rely on a sample to draw inferences about liability.³ For another example, statistical models are often used to estimate damages in antitrust and other complex cases where losses

* G. Rollie White Professor of Law, The University of Texas School of Law. A draft of this paper was presented to the Actuarial Litigation Conference at the University of Connecticut School of Law. I am grateful to the conference participants for useful comments and insights and to my anonymous peer reviewer for excellent suggestions that improved the article.

¹ See, e.g., J. THOMAS MCCARTHY, 4 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 23:1 et seq. (4th ed. 1996). Courts use a multi-factor test to infer likelihood of confusion from factors such as the strength of the mark, the similarity of the marks, and the proximity of the products, but any inference is probabilistic and necessarily refers to statistical properties of the relevant consumer population. See *id.* § 23:19.

² See GEORGE RUTHERGLEN, EMPLOYMENT DISCRIMINATION LAW: VISIONS OF EQUALITY IN THEORY AND DOCTRINE 79-89 (3d ed. 2010).

³ See, e.g., MANUAL FOR COMPLEX LITIGATION § 11.493 (4th ed. 2004).

must be measured relative to a counter-factual baseline that cannot be easily reconstructed using non-statistical techniques.⁴

In fact, the law is bound up with statistical generalization at a very deep level. Any general rule reflects statistical generalizations about a large population of regulated phenomena, whether the generalization is done through the use of formal methods or through informal guesses or even rough political compromise. For example, the general rule that drivers must not exceed fifty miles per hour on a stretch of roadway is based on estimates of the average risk of harm at speeds in excess of fifty miles per hour. Thus, when the speed limit is applied to an individual driver, the driver's liability is evaluated not by the risk that she actually created, but rather by the average risk aggregated over all drivers in all possible situations.

This brief account might lead one to conclude that statistical methods fit litigation smoothly.⁵ But as we know, the use of statistics is controversial. Perhaps the most controversial yet important application is the use of sampling to adjudicate mass tort or other large-damage cases by extrapolating from sample outcomes. Whether the extrapolation involves simple averaging or more complex regression techniques, the result is the same. Very often some cases receive outcomes that differ systematically from the outcomes those cases would have received if they had been tried individually.

Despite the problems, however, there are well-known cases in which courts have used sampling to determine damages and sometimes liability as well. In *Hilao v. Estate of Marcos*,⁶ for example, the Ninth Circuit approved the use of sampling to award compensatory damages in 9,541 consolidated cases.⁷ On the advice of a statistical expert, the district judge randomly selected a sample of 137 cases and used the sample cases

⁴ Another example is calculating backpay for class members in an employment discrimination case involving discriminatory hiring and promotion. See *Pettway v. Am. Cast Iron Pipe Co.*, 494 F.2d 211, 258-63 (5th Cir. 1974). The court must somehow imagine what would have happened to each class member had there been no discrimination. This is such a complex polycentric problem that there is no other feasible method to do it than to use statistical models.

⁵ In his paper *Probability Sampling in Litigation* and his presentation to the Actuarial Litigation Conference, Professor Joseph Kadane gave more examples of the use of statistics in litigation.

⁶ *Hilao v. Estate of Marcos*, 103 F.3d 767 (9th Cir. 1996).

⁷ *Id.* at 782.

to generate a total compensatory damage award for the entire group.⁸ Perhaps the most famous sampling case is *Cimino v. Raymark Industries, Inc.*,⁹ in which Judge Parker, plagued by an onslaught of asbestos litigation, employed sampling to determine individual damages in 2,298 consolidated asbestos cases.¹⁰ He constructed a stratified sample of 160 cases, tried the sample cases, and gave the sample mean to all the other cases in the aggregation.¹¹ The Fifth Circuit reversed, holding that Judge Parker's use of sampling infringed the Seventh Amendment jury trial right and impermissibly altered state substantive law in violation of *Erie*'s dictates.¹² But this case still stands as a dramatic reminder of what might still be possible in some circumstances.

Most recently, an en banc Ninth Circuit Court of Appeals, following *Hilao*, gave a favorable nod to the use of sampling to determine back pay for class members in a massive and highly publicized Title VII class action, *Dukes v. Wal-Mart Stores, Inc.*¹³ The court did so over Wal-Mart's objection that Title VII gave it individualized defenses that could only be adjudicated in individual suits.¹⁴ The Supreme Court reversed. The Court objected to the sampling procedure—calling it “Trial by Formula”—on the ground that sampling impaired Wal-Mart's entitlement to “litigate its statutory defenses to individual claims” and thus violated the Rules Enabling Act.¹⁵ Even so, there is still room left for sampling in future cases. It is not clear how far the Court's objection extends, and in any event, its Rules Enabling Act rationale does not apply to sampling that is legislatively authorized.

⁸ See *id.* The district court did not simply apply the sample average. Instead, a special master made damage recommendations for different injury subgroups by relying on the results of discovery in the sample cases. Then a jury heard testimony on the sampling procedure and special master's recommendations with freedom to reject, accept, or modify the results.

⁹ *Cimino v. Raymark Indus., Inc.*, 751 F. Supp. 649 (E.D. Tex. 1990).

¹⁰ *Id.* at 653.

¹¹ *Id.*

¹² *Cimino*, 151 F.3d 297, 320-21.

¹³ *Dukes v. Wal-Mart Stores, Inc.*, 603 F.3d 571, 625-28 (9th Cir. 2010); see *id.* at 627 n.56 (noting that the invalid claim rate for the sample could be applied to the entire aggregation). Any sampling would take place, however, only after the plaintiffs succeeded in proving company-wide discrimination and thus prima facie liability. See *id.* at 643.

¹⁴ See *id.* at 624-25.

¹⁵ *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541, 2561 (2011).

This Article, prepared for the Actuarial Litigation Conference held at the University of Connecticut School of Law, addresses the normative questions raised by these and other controversial uses of sampling.¹⁶ In addressing these questions, it is important to distinguish between an outcome quality metric and a process-based participation metric. An outcome metric focuses on the quality of the judgments and settlements that sampling produces. Evaluation of outcomes in turn depends on whether one takes a utilitarian or a rights-based approach. The utilitarian evaluates outcome quality in terms of aggregate social benefits and costs. The rights-based proponent evaluates outcome quality in terms of how effectively parties' rights are enforced.

By contrast, a process-based evaluation ignores outcome effects altogether and focuses instead on the intrinsic value of participation. According to the United States Supreme Court, each individual has a due process right to her own personal "day in court"; that is, her own opportunity to control litigation that binds her.¹⁷ If this day-in-court right guarantees individual participation in all cases, it poses a serious obstacle to sampling, since sampling imposes outcomes on parties without giving them an opportunity to litigate their own suits. However, the day-in-court right is not absolute. The question then is what reasons for using sampling justify limiting party participation opportunities consistent with a process-based approach.

I explored these normative questions in a previous article, *Statistical Adjudication: Rights, Justice, and Utility in a World of Process Scarcity* (which I shall refer to as *Statistical Adjudication* for short).¹⁸ I summarize the main points of that earlier article here and extend its analysis in four respects. First, I analyze the effect of sampling on settlement and discuss in more detail the problem of frivolous and weak

¹⁶ The Actuarial Litigation Conference was held on April 15, 2011, and it focused on the use of sampling to litigate mass tort cases in the form of class actions or large-scale, non-class aggregations. For another discussion of the normative issues, see Alexandra D. Lahav, *The Case for "Trial by Formula"*, TEX. L. REV. (2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1945514 (arguing that sampling better achieves equal treatment among litigants and greater transparency of outcomes). See also Alexandra D. Lahav, *Rough Justice*, 1, 30 (2010), available at <http://ssrn.com/abstract=1562677>.

¹⁷ See, e.g., *Taylor v. Sturgell*, 553 U.S. 880, 892-93 (2008).

¹⁸ Robert G. Bone, *Statistical Adjudication: Rights, Justice, and Utility in a World of Process Scarcity*, 46 VAND. L. REV. 561 (1993) [hereinafter Bone, *Statistical Adjudication*].

filings. I ignored settlement effects in *Statistical Adjudication* and only touched on the frivolous suit problem. Both points deserve more extensive treatment. Second, I expand in this Article on the implications of an outcome-oriented rights-based theory for sampling. I discussed this topic in *Statistical Adjudication*, but the following discussion generalizes that analysis in a useful way. Third, I offer some further thoughts about process-based participation and the day-in-court right based on my more recent work.

The fourth extension deals with a possible objection to sampling that I did not discuss in *Statistical Adjudication*. This objection is distinct from adverse effects on outcome and limitations on individual participation. Simply put, it insists that sampling is incompatible with what adjudication is supposed to do. I believe that this objection rests ultimately on an assumption that adjudication at its core involves reasoned deliberation that engages the facts of particular cases. The problem with sampling from this perspective is that it substitutes a formulaic method for case-specific and fact-sensitive reasoning.

I shall refer to this objection as the “methodological legitimacy objection” to highlight its focus on legitimacy and its assumption that legitimacy has to do with the method of decision making rather than the quality of outcomes or the degree of participation. In theory, the methodological legitimacy objection retains whatever force it has even if there is no reason to worry about externalities, party participation is adequate, and litigation costs are reduced. In practice, however, it is likely to operate, when it does, behind the scenes, as a factor influencing decisions to reject sampling on other grounds. For example, the *Wal-Mart* Court might have had something like this concern in mind when it went out of its way to characterize sampling pejoratively as “Trial by Formula.”¹⁹ In any event, the methodological legitimacy objection has sufficient plausibility and superficial appeal to warrant separate discussion even if it is difficult to tell when it is being invoked. As we shall see, the objection is very difficult to sustain in a rigorous way.

The body of this Article is divided into four parts. Part I frames the problem more precisely. Part II focuses on outcome effects with special attention to settlement and frivolous and weak lawsuits. Part III focuses on process-based participation and adds some further thoughts on the day-in-court right. Finally, Part IV discusses the methodological legitimacy objection. Throughout, I mean to consider applications of sampling to

¹⁹ See *Dukes*, 131 S. Ct. at 2561.

determine damages, liability, or both. To be sure, there are special problems with applying sampling to determine liability and much of what I say fits damage sampling better than liability sampling. But liability sampling has been done in the past and might be done more often in the future if process scarcity becomes an even more pressing concern.²⁰

II. FRAMING THE PROBLEM MORE PRECISELY

Sampling runs the risk of distorting outcomes relative to individual trials and substantive entitlements, deprives parties of participation opportunities, and calls for a decision-making method that might be at odds with the usual case-specific reasoned deliberation associated with adjudication. On the positive side, sampling saves litigation resources, helps to equalize litigating power across the party line, and improves real recovery for plaintiffs trapped in a lengthy litigation queue. Parts II, III, and IV explore the normative tradeoff.

Before doing so, however, it is important to clarify the precise nature of sampling's effect on outcome and the normative problem sampling creates from an outcome quality perspective. Some commentators claim that sampling produces more accurate outcomes than individual trials in many situations.²¹ The truth, however, is not nearly as rosy as these claims suggest.

To see why, let us compare the result from an individual trial of a tort case with the result for the same case when it is part of a mass tort aggregation subject to sampling. There is, of course, an error risk associated with an individual trial. Suppose the same case is tried over and over again. If the defendant is in fact liable and juries are reasonably reliable, we would expect most, but not all, of the trials to end in plaintiff verdicts. Moreover, the distribution of damage awards, with the incorrect defendant verdicts counted as zero, should roughly resemble a bell-shaped curve (i.e., a normal distribution) with a possible spike at zero. The mean of this distribution will closely approximate the expected trial outcome, and the mean of the distribution without the zero awards will closely

²⁰ For more on sampling to determine liability elements, see Laurens Walker & John Monahan, *Sampling Liability*, 85 VA. L. REV. 329 (1999).

²¹ See, e.g., Michael J. Saks & Peter David Blanck, *Justice Improved: The Unrecognized Benefits of Aggregation and Sampling in the Trial of Mass Torts*, 44 STAN. L. REV. 815, 851 (1992).

approximate true damages.²² In addition, the standard deviation of the distribution – that is, the spread around the mean – measures the error risk from an individual trial. Let us call this distribution the “individualized error distribution,” or IED for short.

Now assume instead that the case is part of an aggregation of 1000 mass tort cases. Suppose 10% (100 cases) are sampled and tried and that the average of the sample verdicts is calculated (with defendant verdicts assigned a value of zero). Imagine that we repeat this process over and over again. Each time we sample 100 cases randomly, try each of the 100 cases, and calculate the sample average. Not all the samples will be the same, of course, and the sample averages for the different rounds will vary a bit. Nevertheless, if we graph all the sample averages for all the rounds, they should form a bell-shaped (normal) distribution. Let us call the distribution of sample averages the “sample average distribution,” or SAD for short.

The mean of SAD with the erroneous zero verdicts closely approximates the average expected trial outcome for all the 1000 cases in the aggregation, and the mean of SAD without the zero verdicts closely approximates the average true damages for all cases. Moreover, according to basic statistical theory, the standard deviation of SAD should be small and it should get smaller as the sample size increases. In other words, the sample averages cluster rather tightly about the mean and they cluster ever more tightly with increasing sample size.²³ This means that sampling gives a very good estimate of damages for the average case. But it also means that sampling gives a rather poor estimate of damages for those cases that deviate substantially from the average.

²² For example, assume that the defendant caused \$100,000 in damages and that the case is tried 100 times. Suppose that there is a 10% risk of error in determining liability, so 90 of the trials yield plaintiff verdicts and 10 yield defendant verdicts. Also, suppose that of the 90 plaintiff verdicts, 25 are for \$50,000, 25 are for \$150,000, and 40 are for \$100,000. The mean of the entire distribution, including the 10 defendant verdicts, each counted as zero, is \$90,000, which is the same as the *expected* trial verdict when the probability of error in determining liability is 0.1 (i.e., likelihood of proving liability (0.9) x the expected damage amount if liability is proved (100,000) = \$90,000). Considering only the distribution of the 90 plaintiff verdicts, the mean is \$100,000, which is the true damage amount for the case.

²³ For a discussion of this and other statistical properties of the sample average, see RICHARD J. LARSEN & MORRIS L. MARX, AN INTRODUCTION TO MATHEMATICAL STATISTICS AND ITS APPLICATIONS (2d ed.1986).

To be more precise, if the cases in the aggregation do not vary much in salient characteristics (i.e., the aggregation is strongly homogenous), then all the cases closely resemble the average case and as a result the sample average is a very good approximation for every case. Moreover, if the error risk associated with an individual trial for each case is high (i.e., the standard deviation of the IED is large enough), then an individual trial does a relatively poor job of accurately determining case outcomes. With sampling doing a good job and individual trials doing a poor job, it is easy to see that sampling can produce a more accurate outcome than an individual trial for each case.

This is what the proponents of sampling are keen to point out—and it is a very important observation, one not clearly understood by judges and lawyers. The problem, however, is that this happy result breaks down when the aggregation is heterogeneous or the error risk associated with an individual trial is relatively small, or both.²⁴ Indeed, it does not take much heterogeneity before the sample average gives an estimate that is inferior to an individual trial for at least one case in the aggregation.²⁵ Whether this is a normative problem depends on one's theory of adjudication. As Part II.B explains, a utilitarian theory can accommodate a good deal of heterogeneity, but a rights-based theory is less forgiving.

The degree of population heterogeneity and the magnitude of the error risk for individual trials are both empirical questions, and there might be reasons to believe that the former is small and the latter large for some case aggregations. However, there are some, and perhaps many, aggregations for which this will not hold true. Even worse, judges will often find it difficult to determine which aggregations meet the

²⁴ For a more detailed explanation and an example, see Bone, *Statistical Adjudication*, *supra* note 18, at 577-87.

²⁵ The intuition is easy to grasp. When we take a sample, we know that the sample average is very likely to be close to the value of the average case for the population as a whole. This follows directly from the statistical property of the SAD mentioned in the text. Consider a case located at an extreme of the distribution of cases in the aggregation. This case will have a value much higher, or much lower, than the average case and thus the sample average. As long as judges and juries do a reasonably good job of deciding cases accurately on average and make only random errors that are not systematically skewed to one side, it follows easily that an individual trial is likely to come closer to the true outcome for the extreme case than the much lower (higher) sample average.

homogeneity and error risk conditions and which do not.²⁶ Assessing the relevant variables requires specific information about how individual cases vary over the aggregation and how much error individual trials create.²⁷ This case-specific information is costly to obtain, and those costs are precisely what sampling is meant to avoid.²⁸

In sum, my point is that the mean of the IED (which is the expected outcome from an individual trial) for at least some cases is likely to differ from the mean of the SAD (which is the expected outcome in the same case if sample averaging is used). The same is true, although in a bit more complicated way, for more sophisticated regression techniques.²⁹

These points distinguish the use of statistical methods to adjudicate case aggregations from the more accepted uses of statistics in litigation mentioned in the Introduction. Using sampling to extrapolate case outcomes from sample cases is not expressly authorized by any substantive law of which I am aware. Nor is it strictly necessary in the same strong way that statistical models are necessary to construct the counterfactual world for determining damages in complex antitrust cases. Calculating damages in an individual tort suit is a much more straightforward process than reconstructing what the market would have looked like without an illegal antitrust conspiracy or unlawful attempt to monopolize.

It is important to be clear about this last point. Sometimes advocates of sampling point to the *impossibility* of adjudicating individual cases for an extremely large population, such as hundreds of thousands of asbestos cases or the more than one million individual Title VII suits that were aggregated in the *Dukes v. Wal-Mart* class action. This way of

²⁶ This point raises an important question. What level of confidence in the degree of homogeneity and the error risk in individual trials should be required before sampling is justified? Suppose a judge is convinced that it is more likely than not that the aggregation is sufficiently homogenous so that the sample average will yield a more accurate result for each case than an individual trial. Should this be enough, or should something less than preponderance suffice?

²⁷ For example, some cases in a mass tort aggregation will be weaker on liability elements than others (such as a smoker who has trouble proving specific causation); some cases will have weaker evidence to support legal requirements for obtaining damages, and some cases will have more serious injuries than others (such as cancer versus benign abnormalities).

²⁸ Stratified sampling can reduce these problems. However, one must still know a good deal about the population of cases to form sufficiently homogenous subgroups for a stratified sample.

²⁹ See Bone, *Statistical Adjudication*, *supra* note 18, at 584-87.

framing the argument has rhetorical force, but it is incorrect. The fact is that individual litigation is not technically impossible.³⁰ Simply set up a queue and proceed. Matters are very different for the complex antitrust case. Calculating damages in a way that bears any reasonable relationship to those actually suffered is *analytically impossible* without using a statistical model. But it is not analytically impossible to litigate each mass tort or Title VII case individually. It might take hundreds of years to do it (although these kinds of numbers are usually hyperbolic given the inevitability of settlement), but in theory it can be done.

This is an important point because it highlights the respect in which sampling can force plaintiffs who have superior litigating advantages, such as better lawyers, better cases, or simply a better position in the litigation queue, to forego those advantages and accept average outcomes significantly less than the actual value of their cases. To be sure, some of these advantages are a matter of luck and not properly the subject of a moral claim. I shall discuss this point later when I examine the rights-based arguments against sampling.³¹ For now, the important point is that the normative issues must be squarely addressed in the sampling context; they cannot be dodged simply by arguing that there is no other way to provide relief to anyone. By contrast, in the antitrust case, no plaintiff can complain that she would have done better without the statistical approach, because the statistical approach is analytically essential to provide her with any meaningful relief at all.

None of this means, of course, that there are no good reasons to use statistical methods to adjudicate mass tort or other large-scale case aggregations. For one thing, individual trials generate unacceptably high costs in a world of scarce judicial resources.³² Moreover, separate trials generate delay costs for plaintiffs late in the litigation queue and those costs

³⁰ The pure epidemiological mass tort suit might be an exception. See Samuel Issacharoff, *Private Claims, Aggregate Rights*, 2009 SUP. CT. REV. 183, 215-20 (using Vioxx as an example of an “epidemiological mass tort” in which individual recovery is impossible because drug use leaves no trace of evidence to prove individualized causation, even though epidemiological studies confirm a correlation between use and injury). Professor Issacharoff argued at the Actuarial Litigation Conference that the only hope for recovery in these cases is to aggregate all the individual suits and use epidemiological statistics to generate an aggregate damage award. For more on this example, see *infra* note 67.

³¹ See *infra* note 67 and accompanying text.

³² This is, of course, due in large part to very restrictive nonparty preclusion rules. See *Taylor v. Sturgell*, 553 U.S. 880, 882-83 (2008).

can substantially erode the real value of any recovery. This dim prospect creates strong pressure to settle early and on terms favorable to the defendant. Sampling removes this type of unfairness. In addition, statistical methods facilitate aggregate litigation, which helps to equalize litigating power across the party line and produce settlements and trial outcomes closer to the substantive law ideal.³³

These are very weighty reasons. The question, however, is whether reasons like these can justify imposing on some parties statistically generated outcomes that are likely to deviate systematically from their substantive entitlements and from the results of individual trials. In analyzing this question, one should distinguish between consensual and nonconsensual use of sampling and between use to extrapolate final judgments and use to facilitate voluntary settlements.³⁴

This Article focuses on *nonconsensual* sampling used to impose *final judgments*. This is the most controversial application because it is supported neither by consent to sampling itself nor by consent to the settlements that sampling facilitates. Thus, it is the most difficult to justify. Moreover, it is also the most important application. It turns out that justifying nonconsensual use is critical to justifying sampling more generally because many of the uses that seem consensual are on closer inspection less consensual than they first appear.

Let me explain this last point a bit more clearly. In large case aggregations, individual plaintiffs are not likely to be the ones who give consent. The attorney usually decides whether to agree to sampling and whether to settle, and in a world of high agency costs typical of mass tort aggregations, attorneys cannot always be trusted to represent the interests

³³ Roughly, by aggregating separate claims into a single lawsuit, plaintiffs achieve economies of scale and incentivize their attorney to invest more than she would in an individual suit and at a level that is closer to what the defendant is likely to invest.

³⁴ Judges sometimes try a sample of cases from a large aggregation not to impose final judgments, but rather to generate a common baseline of trial verdicts from which parties can estimate the value of their own cases for settlement purposes. Because parties use the sample verdicts as a common baseline, their respective valuations are likely to converge, which makes settlement more likely. Moreover, the randomness of the sample helps to reduce the variance of party estimates, and the judge can reduce variance even further by increasing the sample size (although this also increases costs). For a useful discussion of the benefits of sampling to facilitate settlement in large case aggregations, see Alexandra D. Lahav, *Bellwether Trials*, 76 GEO. WASH. L. REV. 576 (2008).

of plaintiffs faithfully.³⁵ Given that lawyer-client incentives diverge, one should question the extent to which party consent actually legitimates sampling.

There is another reason to question consent in the sampling context. Any consent is likely to be thin. To see why, start with the premise that a party will agree to sampling whenever she expects a better outcome from sampling than from an individual trial. It follows that if delay costs are high enough with a long litigation queue so that the expected value of a trial outcome in the plaintiff's case is virtually zero, a plaintiff should be willing to accept virtually any kind of sampling procedure. But then consent is not meaningful because the plaintiff's choices are radically limited. Sampling might still be justified—and I shall argue in Parts II and III that it is—but it must be justified without relying on consent.

Thus, it is critical to justify nonconsensual use of sampling. As the basis for consent weakens, the need for an independent justification grows stronger, and any independent justification of sampling must include nonconsensual use. There are also other reasons to put nonconsensual use center stage. Parties are not always able to settle even with the benefit of a judicially created baseline, and the settlement process creates transaction costs that could be avoided if the judge were simply to give all parties the average or regression result. So there are efficiency advantages to coercive imposition as well.

III. AN OUTCOME-BASED ANALYSIS

Thus, the question is: When and why can courts use sampling to generate final judgments that are imposed on parties without their consent? The following discussion analyzes this question. It first summarizes the likely effects of sampling on trial judgments and settlements, and then reviews the normative arguments from utilitarian and rights-based perspectives.

³⁵ Both class actions and non-class aggregations are plagued by agency problems. See, e.g., Howard M. Erichson, *Informal Aggregation: Procedural and Ethical Implications of Coordination Among Counsel in Related Lawsuits*, 50 DUKE L.J. 381, 464-65 (2000); Bruce Hay & David Rosenberg, "Sweetheart" and "Blackmail" Settlements in Class Actions: Reality and Remedy, 75 NOTRE DAME L. REV. 1377, 1390-91 (2000).

A. LIKELY OUTCOME EFFECTS

1. Litigated Judgments

Statistical Adjudication made three main points about the effect of sampling on litigated judgments.³⁶ First, the sample average can deviate from a trial judgment for at least one and possibly many cases in the aggregation depending on the degree of heterogeneity. This is the same point as the one developed in Part I above. Second, it is possible to reduce this risk by adjusting the sampling procedure and using regression rather than sample averaging, but these refinements require information about the population of individual cases, which is costly to obtain. Third, sample averaging distorts litigation investment incentives by introducing a new source of free rider and externality problems.

As to the third point, the precise nature of the distortion depends on four factors: (1) whether the sample cases receive the sample average or their own trial verdicts; (2) whether the trial costs in the sample cases are spread over all cases in the population or left for the parties in the sample cases to bear; (3) the pattern of multiple representation of plaintiffs, and (4) the severity of agency problems in a large case aggregation with contingency fees.³⁷ Some combinations of these factors skew litigation investment incentives and results in the defendant's favor.³⁸ Other combinations skew incentives and results in the plaintiffs' favor. *Statistical Adjudication* proposed ways to mitigate these adverse effects.³⁹

In short, sampling can alter outcomes relative to litigated judgments in individual trials, and can do so in ways that for some and perhaps many cases deviate systematically from what the parties' substantive entitlements require. But sampling also produces benefits for many parties and for society at large by reducing cost, risk, and delay. How one strikes the balance depends on whether one takes a utilitarian or a rights-based perspective, as Section II.B below explains.

³⁶ See Bone, *Statistical Adjudication*, *supra* note 18, at 576-94.

³⁷ See *id.* at 587-94.

³⁸ Assuming that the party who invests more is more likely to win.

³⁹ Bone, *Statistical Adjudication*, *supra* note 18, at 587-94.

2. Settlements

The effect of sampling on settlement incentives is complicated. To begin with, it depends on the sampling protocol and in particular on three aspects: (1) whether the sampled cases are allowed to settle after they are chosen for the sample; (2) whether the sample plaintiffs receive the sample average or their own trial verdict; and (3) whether trial costs are averaged and spread over all cases in the population or left for sample plaintiffs to bear.

First, consider the question whether sample cases should be allowed to settle. On the one hand, forcing trial without party consent seems problematic. On the other hand, the point of sampling is to generate trial outcomes from which to extrapolate, and allowing parties in the sampled cases to settle makes it more difficult to achieve this goal. One might simply add settlements to the sample mix, but doing so complicates the task of extrapolation. Settlements are difficult to compare to trial verdicts because settlements discount for likely trial success and are influenced by relative bargaining power, which may or may not correlate strongly with the relative litigating power that affects trial verdicts. To make settlements comparable to trial verdicts, therefore, each settlement must be adjusted to take account of these differences, which is bound to be a complicated and imprecise task.⁴⁰

Given these problems, one might be tempted to exclude settlements, but doing so creates a different set of problems. The cases that settle are not randomly selected, so excluding settlements will taint the randomness of the remaining trial verdicts. Worse yet, it gives the defendant an incentive to settle the strongest sample cases in order to

⁴⁰ Suppose the plaintiff's probability of success in proving liability at trial is p ; the likely damage award conditional on success is w , and the cost to the plaintiff (defendant) of litigating through trial is C_P (C_D). Also assume that the plaintiff's relative bargaining power is γ , meaning that the plaintiff is likely to capture a fraction of the settlement surplus equal to γ . The lowest amount the plaintiff will accept in settlement is $pw - C_P$, and the largest amount the defendant will offer is $pw + C_D$. Therefore, the settlement surplus is $C_P + C_D$ and the likely settlement is: $pw - C_P + \gamma(C_P + C_D)$. If this sample case went to trial and the plaintiff succeeded in proving liability, we would expect a jury verdict close to w . If the cases in the sample vary by w and p , it will be difficult to adjust a settlement of $pw - C_P + \gamma(C_P + C_D)$ so that it is commensurable with verdicts of w in the sample cases that go to trial.

reduce the sample average and thus reduce total liability for all cases in the larger population.

These problems might not be all that serious if a single attorney represents the entire aggregation of plaintiffs and consults her own interest in a fee when she makes the decision whether to settle. Under these circumstances, the attorney has an incentive to counter the defendant's strategy by rejecting settlement offers in the cherry-picked cases. This is so because an attorney who settles cherry-picked cases loses the fee she would have earned with a larger sample average applied to the whole aggregation.⁴¹

Let us assume that the sample cases do not settle, either because settlement is barred or because the attorney rejects every settlement offer. What are the parties likely to do before cases are sampled if they know sampling will be used? Party incentives depend on the other two features of the sampling protocol: whether sample plaintiffs receive the sample average or their own trial verdict, and whether trial costs are averaged and spread over all cases in the population or left for sample plaintiffs to bear. These two elements create four possible scenarios:

	All Cases Receive Sample Average	Sampled Plaintiffs Receive Own Trial Verdicts
Total Costs Spread Over All Cases	SCENARIO I	SCENARIO II
Each Sample Case Bears Its Own Costs	SCENARIO III	SCENARIO IV

Scenario I is attractive on fairness grounds because it treats all plaintiffs in the aggregation equally. But Scenario I might be difficult to implement constitutionally if there are due process problems with denying the parties in the sample cases the benefit of their own trial verdicts. This pushes in the direction of Scenario II. Scenarios III and IV also have some

⁴¹ More precisely, the defendant would have to offer a premium that compensates the attorney for the fee amount lost due to a lower sample average. This is certainly possible but rather unlikely for large case aggregations. Of course, aggregate attorney representation can exacerbate the problem of agency costs, but that is a problem that exists without sampling as well.

attractive features, but there is a strong fairness reason to distribute costs equally over all cases since all plaintiffs in the aggregation benefit from the sample plaintiffs' litigation efforts, and this consideration weighs against Scenarios III and IV. For this reason, the following discussion focuses on Scenarios I and II.

In the Appendix, I present a simple settlement model and use it to analyze the settlement effects of sampling under Scenario I and Scenario II, comparing the results to the no-sampling baseline.⁴² In doing so, I consider two different allocations of settlement power: one in which each plaintiff controls the settlement decision in her own individual case, and one in which all plaintiffs are represented on contingency by the same attorney who controls the settlement decision and settles en-masse.

The most important result of this analysis is that the use of sampling under Scenario I and Scenario II makes settlement impossible, or at least more difficult, for many cases that could have settled without sampling. It follows then that sampling is likely to reduce settlement frequency. Moreover, this is true whether the individual plaintiff or the attorney for the aggregation controls the settlement decision—although the distribution and magnitude of the effects differ. The result for expected settlement amounts is less surprising. High value claims settle for less than they would without sampling. Low value claims normally settle for more.

The following provides a bit more discussion of these settlement impacts, but the details are in the Appendix.

a. Scenario I

The intuition behind the results for Scenario I is easy to grasp. In the simple settlement model, parties can settle if and only if the defendant's expected loss from going to trial is greater than or equal to the plaintiff's expected gain. Expected loss and expected gain depend on each party's estimate of plaintiff's likely success, the expected recovery if plaintiff succeeds, and expected litigation costs through trial. When parties settle, they bargain over how to apportion the savings in trial costs between them, and when they have different estimates of likely success, they also bargain over how to split the additional gains from trade.

⁴² The analysis uses the standard economic model of settlement under asymmetric estimates of likelihood of success. *See, e.g.*, ROBERT G. BONE, CIVIL PROCEDURE: THE ECONOMICS OF CIVIL PROCEDURE (2003) [hereinafter BONE, CIVIL PROCEDURE].

To see this clearly, suppose the plaintiff and the defendant make different estimates of plaintiff's likely success. Let p_π and p_Δ be plaintiff's and defendant's estimates, respectively.⁴³ Suppose the two parties agree on w , the likely recovery if plaintiff succeeds, and on C_P and C_D , the plaintiff's and defendant's costs, respectively, of litigating through trial. These latter two assumptions are not entirely realistic, but they are useful for simplifying the discussion and conveying the basic intuition.⁴⁴

Given that both parties agree on w , the conventional settlement model holds that settlement is feasible without sampling if and only if:

$$p_\Delta w + C_D \geq p_\pi w - C_P$$

The settlement surplus that parties create by settling is the difference between the left hand and right hand sides of this inequality, which is $p_\Delta w + C_D - (p_\pi w - C_P) = (p_\Delta - p_\pi)w + C_P + C_D$. Another way to state the feasibility condition is that the settlement surplus must be greater than or equal to zero; in other words, there must be something for the parties to bargain over:

$$(p_\Delta - p_\pi)w + C_P + C_D \geq 0 \quad (1)$$

Scenario I sampling has two effects on Expression (1). First, it gives all the plaintiffs in the aggregation the sample average for their individual cases. When the aggregation encompasses claims with different valuations, this effect reduces the value of w , the expected recovery conditional on success, for above-average claims and increases it for below-average claims. Second, sampling reduces total litigation costs (i.e., $C_P + C_D$), since only the sampled case are tried. Before the sample is selected, there is a chance that any case could be chosen for the sample, so

⁴³ Thus, p_π and p_Δ might refer to the likelihood of success in establishing liability and proving damages, or they might refer only to the likelihood of success in proving damages conditional on the plaintiff establishing liability without sampling. The referent for the variables depends on whether sampling is used only to determine damages or also to determine liability.

⁴⁴ For example, in the typical case, the plaintiff is likely to be better informed about the seriousness of her injuries (and thus about w) than the defendant. Also, C_P and C_D might vary with case value. It is possible to modify the model to take account of these factors, but doing so complicates the analysis. In a later footnote, I make a few comments about how asymmetric estimates of w might affect the results. See *infra* note 47.

the parties in every case discount litigation costs by the fraction of cases to be sampled. The impact of these two effects on the likelihood of settlement depends on whether $p_\Delta > p_\pi$ or $p_\Delta < p_\pi$.

First, consider the case where $p_\Delta > p_\pi$. In this situation, all the terms in Expression (1) are positive, which means that settlement is feasible for all cases whether or not sampling is used. However, sampling reduces the settlement surplus for above-average cases because w and $C_p + C_D$ both get smaller. For below-average cases, w increases with sampling but $C_p + C_D$ decreases. Therefore, the settlement surplus rises or falls depending on which factor dominates.

Although settlement is feasible in all these cases, the likelihood that parties will reach a settlement can be affected by the size of the settlement surplus. As I explain in the Appendix, one theory holds that parties have greater difficulty reaching a settlement when the settlement surplus is smaller because there is a more limited range of allocations on which the parties can agree. Another theory holds that parties have greater difficulty reaching a settlement when the surplus is larger because they are more likely to bargain hard when more is at stake. Therefore, the effect on settlement depends on which theory of bargaining behavior holds true, which might vary with the circumstances.⁴⁵

Next consider the case where $p_\Delta < p_\pi$. The results here are more striking. If $p_\Delta < p_\pi$, the difference $p_\Delta - p_\pi$ is always negative, so the $(p_\Delta - p_\pi)w$ term in Expression (1) is always negative. Therefore, if w increases enough with sampling (so the negative $(p_\Delta - p_\pi)w$ term gets sufficiently larger in the negative direction) or if $C_p + C_D$ decreases enough with sampling (so the positive term gets sufficiently smaller), a case that has a positive settlement surplus—and therefore could settle without sampling—can have a negative settlement surplus with sampling and be impossible to settle.⁴⁶

⁴⁵ I tend to think that hard bargaining kicks in only for very large settlement surpluses. If I am correct, then we would expect a reduced surplus to make settlement more difficult, unless the surplus is very large both before and after the change.

⁴⁶ The effects vary between above-average and below-average claims in the aggregation. For above-average cases, w decreases with sampling. This means that the negative term $(p_\Delta - p_\pi)w$ is smaller in the negative direction and thus has a weaker impact in reducing the settlement surplus. Still, the magnitude of the reduction in $C_p + C_D$, which depends on the fraction of cases sampled, can be so large that Expression (1) turns from positive without sampling to negative with sampling for above-average cases that are not too far out on the tail of the

These effects obtain whether plaintiffs control their own settlements or an attorney for the aggregation controls the settlement and settles en masse. In the latter case, switching to sampling does not affect w because aggregations settle anyway for average recovery per case multiplied over all cases, which is exactly the same as the estimate under sampling. However, sampling reduces expected litigation costs since only sample cases are litigated, which reduces $C_p + C_D$.

The Appendix develops the analysis more rigorously and describes the different effects that Scenario I sampling can have on the settlement surplus for different types of cases and different sample sizes. The conclusion is the same throughout. For the most likely aggregations, Scenario I sampling rarely, if ever, converts a case that cannot settle into one that can, but frequently converts cases that can settle into ones that cannot.⁴⁷

population distribution. In theory, it is also possible for sampling to turn some cases that cannot settle without sampling into cases that can settle with sampling. However, the Appendix shows that the conditions necessary for this to occur should rarely hold as a practical matter. For below-average cases, w increases with sampling. This means that the negative term gets larger in the negative direction and has a stronger impact in reducing the settlement surplus. This result, combined with the reduction in $C_p + C_D$, guarantees that many below-average cases that could have settled without sampling become impossible to settle with sampling.

⁴⁷ The results are slightly different if the parties have different estimates of w . In the most extreme case, the plaintiff knows w , but the defendant knows only the background distribution of w for all cases in the aggregation (i.e., what fraction are high value and what fraction are low value). Under these circumstances, the defendant must use the average value of w over all the cases; let's denote the average by v . Instead of (1), the settlement condition without sampling for this situation is:

$$p_\Delta v - p_\pi w + C_p + C_D \geq 0$$

For above-average claims, w is greater than v , so it is possible that this condition will not be satisfied when $p_\Delta > p_\pi$ and $w - v$ is very large, in which case settlement is impossible without sampling. (When $p_\Delta > p_\pi$, the condition is always satisfied for below-average claims, i.e. those for which $v > w$.) If Scenario I sampling is used, however, all cases can settle because the plaintiff calculates expected value based on v , the sample average, the same as the defendant does. This means that for above average cases that are located very far out on the tail of the distribution, i.e., where $w - v$ is large enough, settlement can become feasible with sampling when it is impossible without sampling. However, these should be fairly rare occurrences because not many cases are likely to deviate sufficiently from the mean to make this possible. Also, for a very high value claim, the

b. Scenario II

Not much more need be said about Scenario II. It has the same effect on $C_P + C_D$ as Scenario I because litigation costs are shared equally just as they are in Scenario I. But Scenario II sampling reduces the impact on w . This is because sample plaintiffs get their own trial verdicts. All the parties anticipate this possibility because all of them know there is a chance their case will be chosen for the sample, and therefore they include the possibility in their estimates of case value before a sample is chosen. This means that sampling produces a smaller reduction in w for above-average cases and a smaller increase in w for below-average cases. How much smaller depends on the fraction of cases chosen for the sample: the larger the fraction, the smaller the effect.

Nevertheless, Scenario II sampling has the same effect as Scenario I on the two critical factors defining the settlement surplus: it reduces (increases) w for high-value (low-value) claims, and it reduces total litigation costs. This means that it has the same general impact on the likelihood of settlement, except that the ranges of $(p_\Delta - p_\pi)$ values corresponding to the different effects vary to some extent from Scenario I. The precise results are in the Appendix.

In sum, the use of sampling can significantly reduce the settlement rate and thus increase litigation costs, all other things held equal. As the following section explains, this effect is important because it reduces sampling's cost-saving benefits and to that extent weakens the efficiency case for using it. Sampling also gives plaintiffs average recovery, which in effect transfers wealth from high-end to low-end plaintiffs. While this transfer must be justified under both utilitarian and rights-based theories, it is much more problematic for a rights-based theory. The following discussion explores these points.

settlement produced by sampling (which is based on v) departs markedly from the plaintiff's substantive entitlement, which can raise particularly serious fairness concerns.

When $p_\Delta < p_\pi$, the results are also similar to those for the symmetric information case, although the relevant ranges of $p_\pi - p_\Delta$ are different. It is still unlikely that sampling will enable settlement for above average claims, but it is somewhat more likely than in the symmetric information case. Also, sampling never enables settlement and sometimes scuttles settlement for below average cases.

B. AN OUTCOME-ORIENTED NORMATIVE ANALYSIS OF SAMPLING IN LIGHT OF ITS OUTCOME EFFECTS

1. Within a Utilitarian Metric

From a utilitarian perspective, the goal is to maximize aggregate utility, or in the version of utilitarianism associated with law and economics, the goal is to minimize social costs. The social costs of procedure include expected error costs and expected process (or administrative) costs. Thus, procedure aims on this view to minimize the sum of expected error and process costs; i.e., to produce more accurate outcomes but not at the price of excessively costly implementation.

More precisely, the social cost of erroneous outcomes is measured in terms of the policies that the substantive law aims to achieve. An error weakens deterrence and thus distorts primary incentives relative to the substantive law ideal. Process costs include the costs of such things as preparing and filing motions, litigating the issues, holding hearings, and deliberating on a decision. From a law-and-economics perspective, a procedure that reduces error risk might require such a large resource investment that the additional process costs outweigh the marginal reduction in error costs.⁴⁸

In *Statistical Adjudication*, I discussed the efficiency case for sampling.⁴⁹ Extrapolating from the sample average makes a great deal of sense on efficiency grounds. First, as long as aggregations are limited to transactionally-related cases, the sample average should do a reasonably good job of inducing efficient incentives. Agents shape their primary conduct in light of expectations, and the sample average is just an expectation measure. Second, insofar as sampling reduces the delay costs that dilute the real value of a damages payment, it should enhance deterrence. Third, using the sample average can reduce the variance

⁴⁸ To complicate matters further, there are two types of error, false negatives (for example, holding an innocent defendant liable) and false positives (for example, exonerating a guilty defendant). See BONE, CIVIL PROCEDURE, *supra* note 42, at 128-32 (explaining the importance of considering these two types of error). If false negatives are more costly than false positives, a rule might reduce the error risk overall and still increase expected error costs if it reduces the less costly type of error and increases the more costly one.

⁴⁹ See Bone, *Statistical Adjudication*, *supra* note 18, at 595-98.

associated with the expected outcome and thus improve incentives for risk-averse defendants.⁵⁰

On the other side of the coin, sampling adds costs of its own. First, the sampling procedure must be implemented—the sampling protocol designed, the cases actually sampled, and the results analyzed—and this adds process costs. Nevertheless, these costs should be relatively small compared to the litigation and trial costs that sampling saves. Second, by speeding up recovery and attracting more lawsuits, sampling could lead to over-deterrence in some cases. For example, the prospect of having to pay claims sooner could create serious cash flow problems for defendants faced with massive potential liability, and this in turn could force otherwise viable and productive companies into bankruptcy. However, as I argued in *Statistical Adjudication*, these concerns are better handled in ways other than delaying the payment of valid claims.⁵¹ Third, sampling can skew litigation incentives across the party line, and skewed incentives are likely to lead to skewed outcomes. However, the asymmetric stakes in ordinary litigation already produce a skewing effect, and the problems sampling creates can be mitigated to some extent by choosing the right sampling protocol. Moreover, the adverse effects might be offset somewhat if the case aggregation made possible by sampling corrects for a litigating power imbalance across the party line.⁵²

I concluded in *Statistical Adjudication* that the litigation cost savings and beneficial incentive effects make a powerful case for sampling from an efficiency perspective. Moreover, in order to minimize the risk of skewed litigation investment incentives, I recommended that courts use a sampling procedure that gives all plaintiffs the sample average and spreads litigation costs evenly over the aggregation. The following discussion extends this analysis by considering effects on settlement and filing incentives more carefully.⁵³

⁵⁰ This is so when the standard deviation of the distribution of possible sample averages, i.e., the SAD, is less than the standard deviation of the distribution of possible trial verdicts, i.e., the IED.

⁵¹ Bone, *Statistical Adjudication*, *supra* note 18, at 596.

⁵² For a discussion of how aggregation reduces skewed litigation investment incentives, see David Rosenberg, *Mass Tort Class Actions: What Defendants Have and Plaintiffs Don't*, 37 HARV. J. ON LEGIS. 393 (2000); Note, *Locating Investment Asymmetries and Optimal Deterrence in the Mass Tort Class Action*, 117 HARV. L. REV. 2665 (2004).

⁵³ I touched on the filing issue in *Statistical Adjudication*, but I gave it only cursory attention. See Bone, *Statistical Adjudication*, *supra* note 18, at 593-94.

First consider settlement. Section A above showed that a switch from individual litigation to sampling is likely to reduce the settlement rate for cases in the aggregation. With a reduced settlement rate, some (perhaps many) cases that would have settled instead incur additional litigation costs and these costs reduce sampling's cost-saving benefits. The magnitude of this effect, however, is uncertain: it depends in part on when sampling takes place and how much individual litigation precedes it. Sampling's adverse effect on settlement can increase costs substantially when the parties must have ample opportunity to invest in litigation of their individual suits between the time they become aware that sampling will take place and the time that the court actually draws the sample. This pre-sampling investment, after all, is a large part of what is saved by an early settlement. The trial judge can control these costs to some extent by managing the litigation to minimize pre-sampling expenditures and by implementing the sampling protocol expeditiously.

Assuming, however, that there is sufficient opportunity before sampling for parties to invest substantially, the adverse effect of sampling on the settlement rate is likely to be significant and should be included in an efficiency analysis.⁵⁴ The total cost of scuttled settlements increases with the size of the aggregation, so larger aggregations will generate higher costs. Of course, the total cost savings from sampling increase as well. Although it seems reasonable to suppose that cost savings will dominate most of the time, it depends on the fraction of cases that would have settled without sampling and the amount of extra investment those cases incur with sampling.

Second consider frivolous and weak suits. Since only the sample cases are tried, undesirable suits can receive the sample average simply by hiding in the aggregation. One might try to deter this strategy by entertaining summary judgment motions in individual suits before sampling, but doing so would increase pre-sampling costs and magnify the adverse settlement effects discussed in the previous paragraph. It is important to bear in mind, however, that the problem of frivolous and weak

⁵⁴ So too should the effect on settlement quality. The Appendix derives the likely settlement amount assuming equal bargaining power. This should be compared to the expected trial award in individual litigation, assuming that the expected trial award is the proper baseline for assessing deterrence and compensation gains.

filings already exists in the absence of sampling.⁵⁵ Large case aggregations tend to settle en-masse and the attorney for the aggregation has an incentive to include frivolous and weak cases in order to inflate the population size and the ultimate settlement. In the end, it is unclear whether or how much sampling exacerbates these already existing problems.⁵⁶

Even if the problems are more serious with sampling, the use of a sampling procedure makes possible new approaches to managing the risk. For example, a judge might refuse to apply the sample average when the sampling procedure yields a large enough fraction of zero or very small sample verdicts. The idea is to deprive frivolous and weak suits of their cover when the sample results signal a serious frivolous suit problem. This approach wastes the process costs invested in sampling whenever the judge refuses to extrapolate, but it could still make sense if it deterred enough frivolous and weak suits. The important point is that sampling can open up new ways to handle the frivolous suit problem.⁵⁷

2. Within a Rights-Based Metric

The analysis is much more complicated and the conclusions more qualified within a rights-based theory and this is one of the chief reasons sampling is so controversial. In *Statistical Adjudication*, I examined two

⁵⁵ See S. Todd Brown, *Specious Claims and Global Settlement*, 42 U. MEM. L. REV. (forthcoming 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1783792; Francis E. McGovern, *Resolving Mature Mass Tort Litigation*, 69 B.U. L. REV. 659, 688 (1989).

⁵⁶ The defendant who anticipates this strategy can try to counter it by offering a smaller aggregate settlement or even refusing to settle outright. One might think that this is an important difference from sampling, which does not give the defendant this type of control. However, the defendant's total liability with sampling is not affected by frivolous and weak suits because the sample average takes account of their presence. It is the meritorious plaintiffs who are hurt, since they receive a sample average diluted by the presence of frivolous and weak suits in the sample mix.

⁵⁷ The literature on statistical techniques for sorting fraudulent from legitimate insurance claims might provide useful insights. See generally Richard A. Derrig, *Insurance Fraud*, 69 J. RISK & INS. 271 (2002) (providing an overview); Patrick L. Brockett, Richard A. Derrig, Linda L. Golden, Arnold Levine & Mark Alpert, *Fraud Classification Using Principal Component Analysis of RIDITs*, 69 J. RISK & INS. 341 (2002) (proposing a statistical technique for sorting claims).

versions of an outcome-oriented, rights-based theory.⁵⁸ One version assumes that legal rights are designed to enforce moral rights. In other words, it looks through the legal right to focus on the moral right that the legal right protects. The other version assumes that legal rights have force as utility-checking rights independent of their underlying justifications. Accordingly, it focuses on the positive legal right that the substantive law creates.

The existence of substantive rights, whether moral or legal in character, necessarily implies the existence of procedural rights. For without procedural rights, substantive rights could be sacrificed on utilitarian grounds—contrary to their status as *rights*—simply by denying the socially costly procedures needed to enforce them.⁵⁹

The core problem for sampling is the same no matter which version of a rights-based theory one adopts. Sampling can produce outcomes for at least some cases that systematically diverge from what moral or legal rights guarantee. This divergence can be justified in a utilitarian theory by relying on the social costs that sampling saves. But this type of justification is not available in a rights-based theory, or at least not available in quite as straightforward a way. A right is supposed to guarantee its holder the treatment it specifies even when the social costs of doing so are high. Thus, it would seem that sampling, by sacrificing substantive rights to achieve social gains, is just what an outcome-based procedural right is meant to prevent.

At first glance, this problem might seem intractable. However, *Statistical Adjudication* explored several ways to address it. In general, there are two possible approaches to addressing the problem. One approach assumes that the use of sampling is a *prima facie* violation of procedural rights, but that the violation is justified when sampling helps to prevent seriously unfair results produced by high litigation costs and protracted delay.⁶⁰ The second approach denies that there is even a *prima facie* violation. It argues that a proper understanding of the rights at stake shows

⁵⁸ Bone, *Statistical Adjudication*, *supra* note 18, at 605-17.

⁵⁹ For an excellent discussion of this point, see RONALD DWORKIN, *Principle, Policy, Procedure*, in *A MATTER OF PRINCIPLE* 72, 93-94 (1985).

⁶⁰ In this approach, the statistical method used must treat all plaintiffs with equal concern and respect and must aim for outcomes that take account of case-specific facts to the extent practically feasible under the circumstances. See Bone, *Statistical Adjudication*, *supra* note 18, at 615-17. The latter constraint might call for a regression analysis in many situations.

that they can make room for sampling, provided that the sampling procedure is properly designed.⁶¹

More generally—and here I expand on my argument in *Statistical Adjudication* to take account of subsequent work—any sensible conception of outcome-based procedural rights must incorporate four factors that together allow for the use of sampling in appropriate circumstances.⁶² First, a sampling procedure generates aggregate liability that closely approximates what the defendant should pay under the substantive law, and it does so regardless of how the total damages are distributed among plaintiffs. The defendant might insist, as defendants do in these cases, that it has a right to contest liability in each individual case, but there is no obvious outcome-based justification for such a right as a normative matter. After all, the defendant's expected loss is the same in both situations. In fact, its total liability is likely to be more accurately measured with sampling.⁶³

Second, it must matter in some way that a plaintiff who obtains a recovery less than her substantive entitlement makes up for the shortfall with the litigation costs that she saves through sampling. One might object that each plaintiff has a right to the remedy that the substantive law guarantees and that this substantive right does not deduct for litigation costs. On this view, any shortfall in recovery would be a reason by itself to condemn sampling on moral grounds. But this view cannot be correct. If it were, severe delay costs would be irrelevant as well. It would be enough that the plaintiff recovered a formal judgment in the right amount even if she did so many decades after her injury.

The reason litigation cost savings matter is that the substantive rights courts enforce are institutional rights and as such take account of the salient features of the institutions in which they operate, including the

⁶¹ In *Statistical Adjudication*, I focus on the nature of the underlying substantive right. I argue for a corrective justice theory of tort law that recognizes a moral right to compensation only for expected loss. Since the sample average measures expected loss, sampling gives each plaintiff exactly what corrective justice requires. *See id.* at 605-15.

⁶² *See, e.g.,* Robert G. Bone, *Procedure, Participation, Rights*, 90 B.U. L. REV. 1011, 1013-18 (2010) [hereinafter Bone, *Procedure, Participation, Rights*]; Robert G. Bone, *Agreeing to Fair Process: The Problem With Contractarian Theories of Procedural Fairness*, 83 B.U. L. REV. 485, 513-16 (2003) [hereinafter Bone, *Agreeing to Fair Process*].

⁶³ This follows from the statistical property that the sample average is very close to the population average. *See supra* note 23 and accompanying text.

courts that enforce them. Accordingly, they take account of the different ways that rightholders obtain redress within the institution of adjudication, including through the litigation costs they save. This point may seem fairly obvious for legal rights, which after all are created with enforcement in mind. But it is also true for moral rights, although in a less obvious way. Courts do not enforce background moral rights directly; they enforce legal rights that instantiate the moral rights institutionally. And those legal rights, as institutional rights, take account of institutional context, including the litigation costs the institution creates.⁶⁴

The third factor goes to the nature of the procedural right itself. Because outcome error is inevitable and because process costs must matter to the amount of procedure any society provides, outcome-oriented procedural rights are most sensibly defined not as rights to some predefined set of specific procedures, but rather as rights to a fair and just distribution of error risk across cases and litigants.⁶⁵ Understood in this way, procedural rights guarantee that each litigant is treated with equal concern and respect in decisions about how error risk is distributed. This means that the overall error risk can be distributed unequally as long as the reasons for doing so accord equal concern and respect to each individual as a substantive-right-holder. Reasons sounding in social utility, standing alone, are too impersonal to meet this condition. However, reasons that focus on how collective gains benefit each individual personally can qualify.

The fourth factor shifts from the rights litigants possess to the duties they owe one another. This is too complex a subject to provide a detailed analysis here. Let me summarize briefly. The American system of litigation is highly adversarial and parties are given broad freedom to control their own lawsuits. These facts might lead one to conclude that the only duties parties owe one another are duties to refrain from obviously objectionable conduct, such as intentionally filing a frivolous suit or imposing costs for the sole purpose of burdening one's opponent. But a closer examination of actual litigation procedure and practice shows that

⁶⁴ This does not collapse moral rights theory into legal rights theory. In contrast to legal rights, moral rights recognize that background moral principles continue to exert independent force on courts. For example, a court has some freedom to adopt a procedure that better implements a moral right even it also distorts the corresponding legal right to some extent.

⁶⁵ See, e.g., Bone, *Procedure, Participation, Rights*, *supra* note 62, at 1015-18; Bone, *Agreeing to Fair Process*, *supra* note 62, at 513-16. This is not the place to parse the content of this right carefully.

the duties of parties are more robust. Indeed, the fact that procedural rights are institutional rights means that parties owe a general duty of fair regard to one another that is tied to what makes adjudication as an institution work fairly for all litigants.⁶⁶

These four factors taken together can justify sampling on outcome quality grounds in a range of circumstances. The defendant's outcome-oriented rights are fully satisfied by a properly designed sampling procedure. Moreover, the plaintiffs' procedural rights are institutional and as a result take account of institutional context, including the reality of high litigation and delay costs. This opens the door to an argument that procedural rights are satisfied for plaintiffs who end up at least as well off net of litigation costs with sampling as without, a group that includes plaintiffs with cases relatively close to the sample mean. As for those plaintiffs with cases further out on the tail of the distribution, they have procedural rights only to equal concern and respect. This means that they are entitled not to specific procedures or a specific result, but rather to a good reason for the outcome they must bear that respects them as individual rightholders. Moreover, they also owe duties of fair regard to others in the aggregation the same as everyone else.

Still, the fact that parties have procedural rights imposes constraints on when sampling can be used. It is not enough, as it is for a utilitarian approach, that sampling reduces net social costs compared to individual litigation. In a rights-based theory, sampling must be a sensible solution to the problem of high litigation costs and long litigation delays and a solution that fits the fact that parties are rightholders.

For example, suppose cost and delay put some litigants at risk of unfair outcomes due only to the (bad) luck of where they happen to end up in the litigation queue. Because one's place in the queue is a matter of luck and no one can make a moral claim to benefit from this luck, it makes sense to evaluate sampling not *ex post*, after queue position is set, but rather *ex ante*, before any plaintiff knows where she is in the queue. From an *ex ante* perspective, all the plaintiffs face an equal chance of filing late and thus an equal chance of suffering unfair delay. Insofar as sampling

⁶⁶ See Robert G. Bone, *The Puzzling Idea of Adjudicative Representation: Lessons for Aggregate Litigation and Class Actions*, 79 GEO. WASH. L. REV. 577, 619, 623-24 (2011) [hereinafter Bone, *The Puzzling Idea*] (arguing that a duty of fair regard is at work when Rule 19 and Rule 23(b)(1)(B) force some plaintiffs to accept outcomes less than what their substantive entitlements guarantee so that other plaintiffs receive minimally fair recovery).

makes aggregation feasible, it responds to this unfairness for each and every plaintiff. That it does so is a justification for its use that accords equal concern and respect to each plaintiff as an individual rightholder.⁶⁷

Thus, as long as sampling does not distort outcomes for high value plaintiffs by too much,⁶⁸ it can be justified as compatible with outcome-oriented procedural rights.

IV. A PROCESS-ORIENTED ANALYSIS

There are reasons to doubt the coherence of a process-oriented participation right in civil adjudication, but I will not discuss those doubts

⁶⁷ A more extreme example is the epidemiological mass tort. *See supra* note 30. The Vioxx litigation is an example. *See In re Vioxx Prod. Liab. Litig.*, No. 05-3700, 2010 U.S. Dist. LEXIS 64388 (E.D. La. June 29, 2010). Vioxx is a drug prescribed for back pain. After Vioxx was on the market for some time, medical research established a statistically significant link to risk of cardiac abnormalities. However, the cardiac events associated with the use of Vioxx are caused by many other factors as well, and Vioxx leaves no signature trace linking it to the injury. As a result, few plaintiffs can marshal the evidence necessary to prove individual causation by a preponderance of the evidence—even though the epidemiological studies show convincingly that the drug is responsible for a significant fraction of cardiac injuries in the population as a whole.

Given this situation, if suits must proceed individually, many deserving plaintiffs would choose not to sue because the chance of success is too small compared to the cost of litigating. Moreover, many of those who did sue would lose on the causation issue. This would result in potentially serious under-enforcement of tort law, which could impair compensation and deterrence goals. One solution is to aggregate the individual suits into a single class action and use sampling to provide an aggregate damage award for the class as a whole. This solution does not deal with the causation-proof problem, which will still produce an aggregate award significantly below what is optimal, but it does deal with the failure-to-sue problem and thus provides some relief to those injured parties who would not otherwise choose to sue. To deal with the causation-proof problem and provide complete relief that holds the defendant fully accountable, one must use the epidemiological studies to craft an aggregate damage award based on the statistical probability of injury overall. But to do this, one must ignore—or at least skirt—doctrinal obstacles in existing tort law.

⁶⁸ This condition would not be satisfied for case populations that have observable features that strongly indicate high variance. Moreover, it might require the use of regression for some aggregations.

here.⁶⁹ Instead, I shall assume that such a right makes sense and briefly explore its implications, just as I did in *Statistical Adjudication*. The discussion in *Statistical Adjudication* explained why a sampling lottery is a just way to distribute participation opportunities when each litigant has a right to his own day-in-court and budget constraints preclude giving everyone a meaningful individual trial.⁷⁰ It also defined the appropriate scarcity conditions for the use of sampling and explored implications for the choice of sampling methodology. In doing the analysis, however, I accepted, for purposes of argument, the Supreme Court's robust version of the right, the so-called right to a personal day in court that guarantees broad freedom to control strategic choices in individual litigation. With a right defined so broadly, it followed that sampling could be used only in relatively narrow circumstances.⁷¹

I now believe that the best account of the day-in-court right, as that right is reflected in settled features of litigation procedure and practice, is much more limited. The particular version of the right that fits the participation opportunities parties actually enjoy falls far short of the relatively unchecked freedom of strategic choice and party control usually associated with the broad version of the day-in-court right.⁷² For example, a plaintiff can be forced to consolidate her case with hundreds, even thousands, of others under the Multi-District Litigation Act.⁷³ The MDL judge often appoints a litigation committee to control litigation strategy on behalf of the group. The result is that attorneys for most plaintiffs have very little, if any, control over litigation strategy. In effect, plaintiffs are forced to accept a group rather than an individual day in court and they are often forced to do so for reasons that sound in efficiency.⁷⁴ Another example is the (b)(3) class action that binds absent class members to achieve judicial economy gains and often does so without giving those

⁶⁹ See Robert G. Bone, *Rethinking the "Day in Court" Ideal and Nonparty Preclusion*, 67 N.Y.U. L. REV. 193, 279-88 (1992) [hereinafter Bone, *Rethinking*] (explaining the reasons for doubt).

⁷⁰ See Bone, *Statistical Adjudication*, *supra* note 18, at 628-50 (arguing that since sampling distributes participation opportunities by lottery, it is justified whenever a lottery is a just distributional device).

⁷¹ *Id.* at 628-34.

⁷² I have described some of these limitations in a recent article. Bone, *The Puzzling Idea*, *supra* note 66, at 614-24.

⁷³ 28 U.S.C. § 1407 (2008); see Bone, *The Puzzling Idea*, *supra* note 66, at 620-22.

⁷⁴ Bone, *The Puzzling Idea*, *supra* note 66, at 620-22.

absentees realistic opportunities to participate.⁷⁵ To be sure, notice must be sent to class members and absentees have a right to opt out, but they are still bound even if notice fails to reach them and even if they do not understand the notice they receive. Moreover, class representatives and the class attorney represent the interests of the class as a whole, not the individual interests of each class member.⁷⁶

These examples and others like them point to a flexible conception of the day-in-court right, one defined by a balance of considerations relevant to assuring that adjudication works fairly and justly for all litigants.⁷⁷ The fact that the day in court is a *right* still rules out routine reliance on minimizing social costs, but as the MDL and class action examples indicate, it does not rule out social cost arguments altogether.

This flexible and institutional conception allows greater room for sampling. To be sure, the right bars routine use of sampling, just as it bars ordinary utilitarian justifications for its use. At the same time, however, sampling might be reconciled with a process-oriented day-in-court right on broader grounds than avoiding serious unfairness. For example, substantial enough litigation cost savings might justify sampling in the same way judicial economy gains sometimes justify truncated participation in MDL and (b)(3) class actions. In fact, the argument for sampling is stronger in some respects than the argument for the class action on process-oriented participation grounds. Sampling allows more individual participation than the class action, since all litigants make some litigation choices before the sampling procedure is implemented.⁷⁸ Also, sampling can be designed to guarantee even more participation, although doing so increases costs. For example, each party in the larger aggregation might be given a chance to object to the sampling protocol before implementation, and perhaps to

⁷⁵ *Id.* at 592-95.

⁷⁶ *Id.*

⁷⁷ For a more extensive discussion, see *id.* at 615-17.

⁷⁸ In fact, there are notable similarities between sampling and the class action. Sample cases usually share many common questions with cases not chosen for the sample. Moreover, the plaintiffs in non-sample cases should be able to point to a case in the sample that is typical of their own, at least if the overall aggregation is not too heterogeneous and the sample is large enough. In addition, there is no reason to believe that the sample cases would not be litigated vigorously or that lawyers litigating those cases would sell out the aggregation, at least no more reason than already exists without sampling.

argue against application of the sample average to her particular case afterward.⁷⁹

It is important to be clear, however, that squaring sampling with process-oriented participation is only one step in justifying its use. As discussed in Part II above, sampling must also pass an outcome-oriented analysis under a utilitarian or rights-based metric. Furthermore, if there is any sense to the methodological legitimacy critique, sampling must be justified separately on legitimacy grounds as well. Part IV addresses this last topic.

V. THE METHODOLOGICAL LEGITIMACY OBJECTION

To set the stage for the legitimacy objection, imagine that the defendant and all the plaintiffs genuinely consent to the use of a sampling procedure and their consent is their own and not just their attorney's.⁸⁰ Also, assume that the sampling procedure is carefully designed to generate a reliable expected outcome for the population of cases as a whole, and suppose too that it significantly reduces litigation costs and does not adversely affect third parties. In other words, sampling in our hypothetical preserves deterrence benefits without harming others and does so at a significantly lower cost than individual litigation. Is there any reason left to object to it?

Many people—and I count myself among them—would answer no. Nevertheless, one has reason to feel a bit uneasy. After all, deciding cases by extrapolating from a sample is a rather strange way to do adjudication. In the traditional ideal, judges focus on the facts of each individual case

⁷⁹ There is one more potential obstacle to sampling: the jury trial right. *See, e.g., Cimino v. Raymark Indus., Inc.*, 151 F.3d 297, 319-21 (5th Cir. 1998) (holding that trial judge's sampling plan violates the defendant's Seventh Amendment jury trial right). Sampling provides jury trials only for the sample cases. Still, if the sample is large enough, each case in the larger aggregation should have at least one case in the sample that is very similar to it and tried to a jury. Moreover, it is not clear that jury trial must be extended to each separate party. After all, the class action binds absent class members without giving them an individual jury trial, and offensive nonmutual issue preclusion can bind a party to a judge decision in a case where that party would otherwise be entitled to a jury trial. *See Parklane Hosiery Co. v. Shore*, 439 U.S. 322, 333-37 (1979). I leave an analysis of the jury trial objection for another occasion.

⁸⁰ Suppose the parties prefer a speedier resolution at a lower cost.

and reason from those facts to a decision for that case.⁸¹ There are exceptions, of course—the class action being the most notable—but the fact that these are exceptions, and some of them rather controversial, tends to prove the general rule.

This uneasiness with sampling might just be a result of unfamiliarity with its use, but I suspect that more is involved. For example, some critics of large-scale aggregation object to procedures like sampling because they believe that aggregative procedure is somehow at odds with what adjudication is about as an institution.⁸² This type of objection might be about adverse effects on outcome-based rights or process-based individual participation, already dealt with in Parts II and III above.⁸³ But it is also possible that the objection runs deeper, that it rests on a view that aggregative procedures like sampling are institutionally incompatible with civil litigation because they force courts to act in ways that are foreign to adjudication.

To illustrate this point, consider the Supreme Court's gratuitous indictment of sampling as "Trial by Formula" in the recent *Wal-Mart* case.⁸⁴ The Court's explicit argument invoked Wal-Mart's supposed entitlement to "litigate its statutory [Title VII] defenses to individual claims", noted that sampling abridges this entitlement, and concluded that sampling violates the Rules Enabling Act for this reason.⁸⁵ This argument

⁸¹ By the traditional ideal, I mean something like Professor Chayes's traditional model of litigation. See Abram Chayes, *The Role of the Judge in Public Law Litigation*, 89 HARV. L. REV. 1281 (1976) (describing two polar models of litigation—traditional and public law).

⁸² See, e.g., MARTIN H. REDISH, *WHOLESALE JUSTICE: CONSTITUTIONAL DEMOCRACY AND THE PROBLEM OF THE CLASS ACTION LAWSUIT* (2009). I also suspect that objections based on the symbolic or expressive value of individual trial fall into this category. See, e.g., Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1391-93 (1971). These objections tend to focus on the institutional benefits of individualized procedure rather than on party rights.

⁸³ For example, Professor Redish invokes the right to individual participation and fits it into a broader theory of democratic legitimacy. He argues, in effect, that many uses of the class action do violence to democratic legitimacy because they deprive class members of the right to individually litigate their own claims, a right that instantiates democratic participation in adjudication. See REDISH, *supra* note 82.

⁸⁴ *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541, 2561 (2011).

⁸⁵ *Id.*

is weak and not well defended in the case.⁸⁶ Moreover, the Court could have made the argument perfectly well without going out of its way to take a gratuitous rhetorical swipe at sampling.

The Court's use of the phrase "Trial by Formula" suggests a strong aversion to sampling on the ground that it substitutes a statistical formula for an individual trial. But why is the use of a formula such a problem? We can only guess at the answer. It is difficult to see how it can be about bad outcomes or about participation rights that the parties would otherwise have exercised. Wal-Mart has no legitimate reason to complain about the outcome. This is because a properly designed sampling procedure will generate a total amount of backpay damages for the class that closely approximates Wal-Mart's aggregate liability—perhaps even more closely than individual trials.⁸⁷ Moreover, although Wal-Mart is not able to litigate its defense to each individual suit with sampling, it does get to participate fully in each sample case.⁸⁸ Furthermore, because each plaintiff probably has too little backpay at stake to justify an individual suit, all plaintiffs share a strong interest in aggregate resolution, which can be accomplished only through some type of aggregate procedure like sampling. So neither outcome quality nor participation rights seem capable of providing an answer to our question. But there is another possibility. Perhaps the Court believes that sampling just does not belong in adjudication because it

⁸⁶ For example, although Title VII recognizes a *substantive* right to individual defenses against backpay awards, it is not clear that the statute also confers a right to litigate those defenses individually. If it does not, it is unclear what "substantive right" is being "abridged . . . or modified" within the meaning of the Rules Enabling Act. 28 U.S.C. § 2072(b) (2006).

⁸⁷ To be sure, the averaging effect of sampling almost certainly will give some class members a smaller, and some a larger, backpay award than their substantive entitlements guarantee, but that does not affect Wal-Mart. It is also worth noting that, while the Court does not refer to the rights of class members, the distribution of backpay among class members can be justified from a rights-based perspective. The typical backpay amount would not support an individual suit and the possibility of qualifying for a (b)(3) class action is remote without the use of sampling. Thus, one can argue that all plaintiffs have an interest in sampling so they can recover at least some backpay award.

⁸⁸ Also, any claim that Wal-Mart might have to process-based participation must take account of the adverse effect on the participation rights of class members, most of whom would probably not be able to bring their backpay claims at all without sampling. Indeed, it is fairly obvious that Wal-Mart seeks individual litigation precisely because it is likely to discourage the pursuit of individual backpay claims.

involves trying cases with a “formula” and not attending individually to the facts of each case.

In any event, I believe that the institutional argument has sufficient intuitive appeal that it should be addressed separately. Even if judges and scholars do not invoke it explicitly, they could still be influenced by it implicitly while formulating their express objections on outcome quality or participation grounds.⁸⁹ In a world of institutional differentiation and specialization, the legitimacy of an institution depends in large part on the presence of structural elements that fit the institution’s distinctive purpose and function.⁹⁰ Following this logic, critics of sampling might say that the function of courts is to decide individual claims of right and that traditional litigation procedure is essential to this function and thus essential as well to the institution’s legitimacy.

I shall refer to this type of argument as the “methodological legitimacy objection” since it focuses on legitimacy and supposes that legitimacy depends on the method used to decide a case. To get a clearer grasp on the nature of the argument, let us consider a more obvious example than sampling. Most people bristle at the idea that a judge would decide an issue by flipping a coin, and they are likely to object even though the issue is in equipoise, each side has an equal chance to win, and no third parties are harmed.⁹¹ The objection is that flipping a coin is simply not a

⁸⁹ In particular, judges might be more willing to embrace an argument that sampling violates outcome-based rights or infringes a litigant’s due process right to a personal day in court because they also believe that sampling is simply not what adjudication is about.

⁹⁰ For example, the legitimacy of the legislative process depends on a voting system that facilitates public participation and, in theory at least, assures representative accountability to electorate preferences. This voting system adds legitimacy because it fits the function of legislation in a way that accommodates democratic values. However, voting would contribute nothing to legitimacy if the legislature were suddenly enlisted to adjudicate individual cases as well. In fact, many would deem it illegitimate for a legislature to take on the function of adjudication, even if the parties agreed and even if all the legislators wanted to do it.

⁹¹ See, e.g., *In re Brown*, 662 N.W.2d 733, 736 (Mich. 2003) (judge censured for flipping a coin when neither side’s argument was more persuasive); Adam M. Samaha, *Randomization in Adjudication*, 51 WM. & MARY L. REV. 1, 28-29 (2009) (providing several examples). The hypothetical assumes that the decision is not subject to a burden of persuasion that would break the tie. To make the situation more concrete, imagine that the issue is committed entirely to the judge’s discretion. It is worth pointing out though that the preponderance-of-the-evidence

proper decision procedure for adjudication. If pressed to explain why, a critic would probably focus on the close link between adjudication and case-specific deliberation.⁹² She might argue that judges are supposed to decide cases by reasoning through the implications of general rules and principles on the facts of the particular case and that this mode of reasoning is essential to adjudication's legitimacy.⁹³

It is important to be clear about the nature of this objection. It has nothing necessarily to do with adverse effects on the substantive or procedural rights of the parties. Moreover, neither coin flipping nor sampling is an arbitrary decision procedure.⁹⁴ A judge can have a very good reason to use either method. For example, flipping a coin can be justified on moral grounds when it is impossible to tell which party is correct and both have equally strong substantive entitlements.⁹⁵ So too, sampling makes sense when the sheer volume of cases produces serious problems for individual litigation, as previously discussed.

The methodological legitimacy objection, I believe, has to do with the fact that sampling, like coin flipping, disables the usual reasoning process at the point of actual decision. The judge relies exclusively on a statistical method rather than applying rules and principles to the facts of each specific case. Still, the question remains why this is an illegitimate method when the judge can provide a sensible reason for using it. The

persuasion burden, as a general rule for breaking ties, is itself based on statistical generalizations about broad categories of cases.

⁹² See Shay Lavie, *Reverse Sampling: Holding Lotteries to Allocate the Proceeds of Small-Claims Class Actions*, 79 GEO. WASH. L. REV. 1065, 1084-85 (2011) (arguing that people oppose lotteries because they substitute luck for reason, and quoting the N.Y. Commission on Judicial Conduct in *In re Friess*, ANNUAL REPORT OF THE NEW YORK STATE COMMISSION ON JUDICIAL CONDUCT 88 (1984), for the proposition that: "The public has every right to expect that a jurist will carefully weigh the matters at issue and . . . render reasoned rulings and decisions.").

⁹³ See generally JON ELSTER, *SOLOMONIC JUDGMENTS* 38 (1989) (emphasizing that "the use of lotteries to resolve decision problems under uncertainty presupposes an unusual willingness to admit the insufficiency of reason.").

⁹⁴ Cf. *id.*, at 102 (noting that randomness in legal decisions is often associated with arbitrariness or whimsy).

⁹⁵ See generally Lewis A. Kornhauser & Lawrence G. Sager, *Just Lotteries*, 27 RATIONALITY & SOCIETY 483, 495-505 (1988) (discussing equal entitlement and scarcity conditions for using the lottery as an exclusive or nonexclusive method of allocation and noting that using the lottery under these conditions is supported by reasons).

answer must be that the application of reason at the point of case-specific decision is a fundamental aspect of adjudication that neither the parties nor the judge can change without risking the institution's legitimacy.

It is quite common to view adjudication as intimately tied to a special reasoning process that combines general principles with case-specific facts.⁹⁶ For example, Lon Fuller characterized common law reasoning in this way. He described a decision process that closely resembles the method of reflective equilibrium.⁹⁷ Roughly, judges interpret the law by placing existing legal principles and norms alongside the facts of the particular case. The judge moves back and forth between her best understanding of the law and whatever moral or practical intuitions the facts generate, adjusting law and intuition until they fit together in reflective equilibrium.

However, even if this account of adjudicative reasoning is correct, as I believe it is, there remains the question why exceptions are not permitted when they respond in a sensible way to serious litigation problems. One possible reason to worry about exceptions has to do with public perception. The concern on this account is that the public will lose faith in the legitimacy of adjudication if judges employ unfamiliar methods to resolve cases. But this concern is exaggerated and ultimately unpersuasive. For one thing, public perceptions are malleable. For example, the public might accept coin flipping as legitimate in a particular case if they knew that the parties requested it and understood that it was supported by good reasons. In addition, public perception is circular. People tend to equate what is legitimate with what is familiar.⁹⁸ If judges routinely flip coins, for example, public opinion could shift toward accepting coin flipping as a proper decision method. Finally, it is simply implausible that the public would give up on the court system just because judges

⁹⁶ See, e.g., RONALD DWORKIN, *LAW'S EMPIRE* 225-75 (1986); Lon L. Fuller, *The Forms and Limits of Adjudication*, 92 HARV. L. REV. 353, 372-81 (1978).

⁹⁷ See Robert G. Bone, *Lon Fuller's Theory of Adjudication and the False Dichotomy Between Dispute Resolution and Public Law Models of Litigation*, 75 B.U. L. REV. 1273 (1995) (describing Fuller's views in terms of reflective equilibrium). On the method of reflective equilibrium more generally, see JOHN RAWLS, *A THEORY OF JUSTICE* 42-45 (rev. ed. 1999).

⁹⁸ The other alternative is to base their opinion on what they believe courts *should do*. But in that case, it is not the perception that matters, but the underlying normative theory that supports the perception.

occasionally used sampling to decide large case aggregations, especially if they also provided good reasons for doing so.⁹⁹

This leaves only one reason I can think of for worrying about exceptions. This has to do with the adverse effect of occasional use on the quality of adjudicative decisions over the long run. The concern is that allowing some exceptions will invite more exceptions and send adjudication down a slippery slope, transforming the institution in undesirable ways. This concern might have force for coin flips. Maybe a few coin flips would not be a problem, but if judges became accustomed to flipping coins, they might relax constraints on its use and make coin flipping a more general practice. Also, a judge faced with a difficult decision might be tempted to give up too soon and resort to flipping a coin when a more careful analysis would show that a reasoned decision is feasible. This could be particularly problematic if hard cases are the ones where principled decision is most valuable for the development of the law.

Whatever merit it might have for coin flips, this slippery slope argument is much less convincing for sampling. No matter how frequently sampling is used, there will always be sample cases decided in the ordinary way. Thus, judges never completely escape individualized decisions. Moreover, there is no reason to believe that *every* case must be adjudicated individually in order to produce good common law rules and principles or sound interpretations of statutes or constitutional provisions. Finally, the use of sampling is limited to large case aggregations and requires much more deliberation and preparation than coin flipping. Thus, the slippery slope is a lot less slippery for sampling than for coin flipping.

In sum, it is not at all clear that the methodological legitimacy objection has force against a well-justified use of sampling in mass tort aggregations. Sampling is sufficiently different from coin flipping even though both employ probabilistic techniques and randomized decision procedures.

⁹⁹ There is a closely related argument that deserves brief mention. According to this argument, adjudication has social value as a symbol of our collective commitment to principled reason in government and that this symbol's message would be diluted if judges flipped coins or used sampling. Even if the premise is true, the conclusion does not necessarily follow. I find it rather far-fetched to believe that the message would be lost if judges sometimes used sampling. Indeed, the fact that sampling is itself supported by good reasons should reinforce the message of reason's importance in government.

VI. CONCLUSION

Sampling is an extremely useful tool for litigating large aggregations of cases. Squaring it with adjudication, however, raises a number of complicated normative questions. In this Article and in my earlier work, I have attempted to address three types of challenges: challenges directed to sampling's effect on outcome quality, challenges directed to its effect on process-based participation, and challenges based on sampling's supposed incompatibility with adjudication's distinctive mode of decisionmaking.

In the end, sampling can be justified in many more situations than courts currently apply it, and society is paying a very high price for ignoring this insight. Courts should be more receptive to the benefits of sampling and judges should engage the task of justifying its use more carefully. The system of adjudication would be much the better for it.

APPENDIX

This Appendix models the settlement decision under a no-sampling regime and under two different sampling scenarios. The point is to show that sampling can often reduce the likelihood of settlement and skew the settlement amount.

The analysis considers settlement incentives before any cases are actually sampled on the assumption that all parties know that sampling will take place and also know the court's sampling protocol. I consider the results when each plaintiff controls her own settlement decision, and then when an attorney representing all plaintiffs in the case aggregation on contingency makes the settlement decision in her own self-interest.

I. MODEL AND TERMINOLOGY

Let N be the total number of cases in the aggregation. Let α be the fraction of cases that will be sampled. So αN is the number of cases in the sample. Assume that each case has a single plaintiff and a single defendant and that the plaintiffs are all different but the defendant is the same. Suppose there are two types of claims in the aggregation, high-value claims (H) and low value claims (L). To simplify the analysis, assume that these two types of claims vary only with respect to the amount of damages and not the objective likelihood of plaintiff's success.¹⁰⁰ Let w^H and w^L be the damages for a high-value and a low-value claim, respectively.

Suppose that the plaintiffs and the defendant know w , but disagree about plaintiff's likelihood of success in proving liability or damages, or both, at trial. This type of disagreement can occur, for example, when there is asymmetric information so that one party has information about the claim not yet known to the other side. Assume all the plaintiffs share the same estimates of likely success, which we shall denote p_π . Let p_Δ be the defendant's estimate of plaintiff's likelihood of success and assume that it is the same for all the cases.¹⁰¹

¹⁰⁰ This is just for purposes of simplification. One can also vary likelihood of success and get similar qualitative results.

¹⁰¹ Therefore, the parties might have different information about liability or they might view generally known evidence of liability differently. Alternatively, they might agree on the probability of liability but disagree on the likely fraction of full damages that the plaintiff will be able to prove. In this case, p can be

Now let β be the fraction of high-value claims in the aggregation. Also, let C_P be each plaintiff's cost of litigating her individual case all the way through trial and let C_D be the defendant's cost.¹⁰² To simplify the analysis, assume that the parties have equal bargaining power in settlement negotiations, so they split the settlement surplus evenly.¹⁰³ Finally, it will be convenient to have a variable to denote the average damage amount over the entire class. Let v be this average, so $v = \beta w^H + (1-\beta)w^L$.

It is worth noting that the assumption that both parties know whether a case is high or low value – and therefore agree on w – is rather strong. It is more realistic to assume that the plaintiff has private information about the value of w for her particular case. Nevertheless, the strong assumption simplifies the analysis and conveys the essential insight. In footnotes, I explain why the results are likely to be similar when information about w is asymmetric.¹⁰⁴

II. THE NO-SAMPLING BASELINE

First, we need to determine the results in a litigation world without sampling. These results will serve as a baseline against which to compare the impact of sampling.

A. PLAINTIFFS CONTROL SETTLEMENT DECISION

Suppose that each plaintiff makes the decision whether to settle and for how much. Without sampling, the conditions for settlement being feasible for a high-value and a low-value claim, respectively, are:

$$\begin{aligned} p_{\Delta}w^H + C_D &\geq p_{\pi}w^H - C_P \\ p_{\Delta}w^L + C_D &\geq p_{\pi}w^L - C_P \end{aligned}$$

interpreted as the probability of success on liability times the fraction of a full damage recovery the plaintiff is likely to receive.

¹⁰² For simplicity, I assume that C_P and C_D are the same for high-value and low-value claims. I could relax this assumption, but it would complicate the analysis unnecessarily.

¹⁰³ We could generalize by letting γ be the plaintiff's relative bargaining power; that is, γ would be the fraction of the settlement surplus that the plaintiff can capture. In this model, I set $\gamma = 0.5$.

¹⁰⁴ See *infra* notes 105-108, 110.

These are just the standard settlement feasibility conditions. The defendant's expected trial loss must be greater than or equal to the plaintiff's expected trial gain for the defendant to be willing to offer a settlement that the plaintiff is willing to accept.

Rearranging, we get:

$$(p_\Delta - p_\pi)w^H + C_P + C_D \geq 0 \quad (1)$$

$$(p_\Delta - p_\pi)w^L + C_P + C_D \geq 0 \quad (2)$$

The expression on the left hand side is the settlement surplus, which must be nonnegative for settlement to be feasible.¹⁰⁵

When bargaining power is equal, as we assume it is, the expected settlement is likely to be at the midpoint of the settlement range. Letting S^{H*} and S^{L*} be the expected settlement for a high-value and a low-value case, respectively, we have:

$$S^{H*} = (p_\pi + p_\Delta)w^H/2 + (C_D - C_P)/2 \quad (3)$$

$$S^{L*} = (p_\pi + p_\Delta)w^L/2 + (C_D - C_P)/2 \quad (4)$$

B. ATTORNEY FOR ALL PLAINTIFFS CONTROLS SETTLEMENT DECISION AND SETTLES EN MASSE

Now assume that all the plaintiffs in the aggregation are represented by the same attorney, who is hired on contingency with a contract that specifies a contingency percentage of r . Suppose that the attorney only settles en masse and that she makes the settlement decision to maximize her own fee; in other words, assume that agency costs are high.

¹⁰⁵ Suppose instead that information about w is asymmetric: the plaintiff knows whether her case is high or low value, but the defendant only knows the background fraction, β , of high value claims. In this situation, the defendant will assign the average value, v , to all cases. Let $z^H = w^H - v$ and $z^L = v - w^L$. Then the conditions for settlement being feasible without sampling, for a high-value and a low-value claim, respectively, are:

$$(p_\Delta - p_\pi)w^H - p_\Delta z^H + C_P + C_D \geq 0$$

$$(p_\Delta - p_\pi)w^L + p_\Delta z^L + C_P + C_D \geq 0$$

Thus, the settlement surplus differs from the symmetric information case by a factor equal to the amount by which the true value of w differs from the average value, discounted by p_Δ .

The smallest settlement the attorney will accept, S , is one that makes her indifferent between settling or going to trial. This condition is: $rS = np_{\pi}v - nC_p$. Thus, the attorney's minimum settlement demand is:

$$np_{\pi}v - nC_p/r$$

The most the defendant is willing to offer is a settlement that makes it indifferent between settling and going to trial. Therefore the defendant's maximum offer for the whole aggregation is:

$$np_{\Delta}v + nC_D$$

The feasibility condition for settlement if settlement takes place en masse is:

$$np_{\Delta}v + nC_D - np_{\pi}v + nC_p/r \geq 0$$

Simplifying, we get:

$$(p_{\Delta} - p_{\pi})v + C_p/r + C_D \geq 0 \quad (5)$$

And S^* for an en masse settlement with attorney control is:

$$S^* = n[(p_{\pi} + p_{\Delta})v + C_D - C_p/r]/2 \quad (6)$$

III. WITH SAMPLING

The parties' expectations change with sampling. A plaintiff knows that if she is chosen for the sample, she will receive either her own trial verdict or the sample average depending on the sampling protocol – and the defendant knows the same thing. If the sample plaintiffs' costs are shared equally by all plaintiffs, then each plaintiff's litigation costs are the same and equal to αC_p . However, if sample plaintiffs must pay their own litigation costs, then the litigation costs for each of the sample plaintiffs are C_p and the litigation costs for each of the remaining plaintiffs are 0.

Let us assume that the defendant in all the scenarios averages total litigation costs for the sampled cases over all the cases in the aggregation. It follows that the defendant's anticipated litigation costs are the same for all cases; namely αC_D .

The following discussion analyzes only Scenarios I and II. The other two scenarios can be analyzed in the same way.

A. SCENARIO I: SAMPLE PLAINTIFFS RECEIVE SAMPLE AVERAGE AND COSTS ARE SHARED EQUALLY

1. When the Plaintiffs Control the Settlement Decision

a. *Effect on Settlement Feasibility*

With these assumptions in place, we can set forth the feasibility conditions for settlement before any sample is chosen. In Scenario I, all the plaintiffs get the sample average and share the sample plaintiffs' litigation costs equally. Therefore, a plaintiff's expected value of litigating through trial when she knows sampling will take place is: $p_\pi v - \alpha C_P$. The defendant's expected loss is: $p_\Delta v + \alpha C_D$. Therefore, the feasibility condition for settlement in Scenario I is:

$$(p_\Delta - p_\pi)v + \alpha(C_P + C_D) \geq 0 \quad (7)$$

Given this, let us examine whether the use of sampling is likely to reduce, increase, or leave unaffected the likelihood of settlement compared with the no-sampling baseline. To determine this, we must compare (7) with (1) and (2). It is useful to consider cases where $p_\Delta \geq p_\pi$ and cases where $p_\Delta < p_\pi$ separately.

First, suppose $p_\Delta \geq p_\pi$. Comparing (1) and (2) with (7), it is easy to see that settlement is feasible for all cases with and without sampling. However, sampling might affect the probability of successful settlement for high value and low value claims. For high value claims, sampling reduces the settlement surplus. This follows directly from the fact that $v < w^H$ and $\alpha < 1$. Whether this is likely to reduce or increase the frequency of settlement depends on how the size of the surplus affects the likelihood of settlement. One view is that a larger surplus creates more points of potential agreement for the parties, which makes settlement more likely. Another view is that a larger surplus invites harder bargaining because there is more to gain, which makes settlement less likely. Under the first view, sampling is likely to reduce the probability of settlement for high-value claims. Under the second view, it is likely to increase the probability.

For low-value claims, the effects depend on the magnitude of $p_\Delta - p_\pi$. In particular, using sampling increases the surplus if $p_\Delta - p_\pi > (1 - \alpha)(C_P + C_D)/(v - w^L)$, which is, after rearranging, $\alpha > 1 - [(p_\Delta - p_\pi)(v - w^L)/(C_P + C_D)]$. For any realistic α , such as a 10% or 15% sample size, this condition is not likely to be satisfied unless $v > w^L$, which

in turn is not likely unless $w^H > w^L$. It follows that using sampling is likely to reduce the settlement surplus for most low-value cases as well.¹⁰⁶

Second, suppose $p_\Delta < p_\pi$. When this condition holds, some cases that can settle without sampling cannot settle with sampling. To see this point, note that the following two conditions must be satisfied if a case can be settled without sampling but not with sampling, if the claim is high value:

$$\begin{aligned}(p_\Delta - p_\pi)w^H + C_P + C_D &\geq 0 \\ (p_\Delta - p_\pi)v + \alpha(C_P + C_D) &< 0\end{aligned}$$

Let $q = p_\pi - p_\Delta$. Solving for q in each inequality and putting the inequalities together, we get:

$$\alpha(C_P + C_D)/v < q \leq (C_P + C_D)/w^H$$

For this to be possible, $\alpha(C_P + C_D)/v < (C_P + C_D)/w^H$, which implies that $\alpha < v/w^H$.

Therefore, for high-value claims with $p_\Delta < p_\pi$ (i.e., $q > 0$), the case can settle without sampling but not with sampling if and only if:

$$\alpha < v/w^H, \text{ and}$$

$$\alpha(C_P + C_D)/v < q \leq (C_P + C_D)/w^H$$

If $q \leq \alpha(C_P + C_D)/v$, then the case can settle with or without sampling, and if $q > (C_P + C_D)/w^H$, then the case cannot settle whether or not sampling is used.

¹⁰⁶ The results are a bit different when information about w is asymmetric. See *supra* note 105. One must compare (7) with $(p_\Delta - p_\pi)w^H - p_\Delta z^H + C_P + C_D$ for high-value claims and with $(p_\Delta - p_\pi)w^L + p_\Delta z^L + C_P + C_D$ for low-value claims. When $p_\Delta > p_\pi$, it is theoretically possible for sampling to enable settlement for high value claims (but never for low value claims) when settlement is not otherwise feasible. For this to hold true for a high value claim, two conditions must be satisfied:

$$\begin{aligned}(p_\Delta - p_\pi)w^H - p_\Delta z^H + C_P + C_D &< 0 \text{ and} \\ (p_\Delta - p_\pi)v + \alpha(C_P + C_D) &\geq 0\end{aligned}$$

The latter condition is always satisfied and the former is satisfied if $z^H > [(p_\Delta - p_\pi)w^H + C_P + C_D]/p_\Delta$. In other words, the case must be quite far out on the tail of the distribution before sampling enables settlement.

The opposite result—i.e., that sampling makes settlement feasible—is also possible but highly unlikely for most aggregations. It can be easily shown that for sampling to enable settlement when it would not otherwise occur, the following condition must be satisfied: $(C_P + C_D)/w^H < q \leq \alpha(C_P + C_D)/v$. This condition can hold only if $\alpha > v/w^H$. But this constraint on α (the sample size) is not likely to hold for most aggregations. As long as the standard deviation of the aggregation is not unusually large, v/w^H will be a reasonably large fraction and no court is likely to sample a large fraction of cases from the aggregation.¹⁰⁷

One can do the same analysis for low-value claims. It is easy to see that the switch to sampling can never make settlement possible for a low-value claim if it is not possible without sampling. This is because $(p_\Delta - p_\pi)v + \alpha(C_P + C_D) < (p_\Delta - p_\pi)w^L + C_P + C_D$ whenever $p_\Delta < p_\pi$ (since $v > w^L$). However, the switch to sampling scuttles settlement for low-value cases whenever $\alpha(C_P + C_D)/v < q \leq (C_P + C_D)/w^L$.¹⁰⁸

To summarize, we have the following two results for cases where $p_\Delta < p_\pi$:

- For realistic values of α and aggregations that are not too widely dispersed about the mean, switching from no-sampling to sampling never turns a case that cannot settle into one that can.
- More importantly, using sampling turns some cases that can settle into ones that cannot. These are cases where $\alpha(C_P + C_D)/v < q \leq (C_P + C_D)/w^i$ ($i = H$ or L).

¹⁰⁷ When information about w is asymmetric, similar results obtain. *See supra* note 105. It is easy to derive the parallel conditions for sampling to scuttle settlement for high value claims, assuming $p_\Delta < p_\pi$:

$$\alpha < [1 - p_\Delta z^H / (C_P + C_D)]v/w^H, \text{ and} \\ \alpha(C_P + C_D)/v < q \leq (C_P + C_D - p_\Delta z^H)/w^H$$

If $\alpha > [1 - p_\Delta z^H / (C_P + C_D)]v/w^H$, there is a range of q for which sampling enables settlement of high value claims, just as for the symmetric information case. However, as long as $p_\Delta z^H / (C_P + C_D)$ is relatively small, α is very unlikely to exceed this threshold and sampling will only scuttle settlement of high-value claims.

¹⁰⁸ Similar results obtain for low-value claims when information about w is asymmetric. Sampling never enables settlement no matter what α is. Moreover, sampling scuttles settlement when $\alpha(C_P + C_D)/v < q \leq (C_P + C_D + p_\Delta z^L)/w^L$.

To give a concrete example of the second result, suppose a high-value claim is worth \$1,000,000 and a low-value claim is worth \$600,000 and 20% of the aggregation is high-value claims. Suppose $C_P = C_D = \$150,000$, and a 10% sample is used, so $\alpha = 0.1$. Then $v = .2 \times 1,000,000 + .8 \times 600,000 = 680,000$, and $v/w^H = 0.68$. Therefore, the condition $\alpha < v/w^H$ is satisfied (and, of course, $\alpha < v/w^L$ for all α , since $v/w^L > 1$). In this case, $\alpha(C_P + C_D)/v = 30,000/680,000 = .044$. For high-value claims, $(C_P + C_D)/w^H = 300,000/1,000,000 = 0.3$. For low-value claims, $(C_P + C_D)/w^L = 300,000/600,000 = 0.5$. Assume $p_\Delta < p_\pi$. If the difference between the plaintiff's and defendant's estimates of p is between 0.044 and 0.3, using sampling will turn all claims into ones that cannot settle.

b. *Effect on Settlement Amount*

Next, consider the effect of sampling on the expected settlement amount. Assuming equal bargaining power, so the parties split the surplus evenly, the expected settlement amount with Scenario #1 is:

$$S^* = [(p_\pi + p_\Delta)v + \alpha(C_D - C_P)]/2 \quad (8)$$

We must compare (8) with (3) and (4). It is easy to see that sampling always reduces the settlement amount of high-value claims – from $[(p_\pi + p_\Delta)w^H + (C_D - C_P)]/2$ to $[(p_\pi + p_\Delta)v + \alpha(C_D - C_P)]/2$. Sampling also increases the expected settlement for low-value claims if $\alpha > 1 - [(p_\pi + p_\Delta)(v - w^L)/(C_D - C_P)]$, which should (almost) always hold true.

2. When the Attorney Controls the Settlement Decision and Settles En-Masse

a. *Effect on Settlement Feasibility*

The condition for a feasible settlement under Scenario I when the attorney is in control is:

$$(p_\Delta - p_\pi)v + \alpha(C_P/r + C_D) \geq 0 \quad (9)$$

We must compare (9) with (5). Doing so yields the following results:

- If $p_\Delta \geq p_\pi$, the aggregation can settle en-masse with and without sampling, but the surplus is less with

sampling. The surplus is $(p_\Delta - p_\pi)v + C_p/r + C_D$ without sampling and $(p_\Delta - p_\pi)v + \alpha(C_p/r + C_D)$ with sampling.

- If $p_\Delta < p_\pi$, then for all α (with $q = p_\pi - p_\Delta$), a case that cannot settle without sampling cannot settle with sampling. But there are cases where settlement is scuttled with sampling. These are cases where $\alpha(C_p/r + C_D)/v < q \leq (C_p/r + C_D)/v$.¹⁰⁹

To illustrate, consider the same example as we analyzed above: $w^H = \$1,000,000$; $w^L = \$600,000$; 20% of the aggregation is high-value claims; $C_p = C_D = \$150,000$, $\alpha = 0.1$, and $v = 680,000$. Assume $r = 0.25$, which is roughly the average contingency recovery in large aggregations. Then $C_p/r + C_D = 750,000$.

If $p_\Delta \geq p_\pi$, then settlement is always possible, but sampling reduces the size of the surplus by \$675,000. This is a significant amount given that v is \$680,000. For example, suppose $p_\Delta - p_\pi = 0.4$. Then the surplus falls from \$1,022,000 to \$347,000.

If $p_\Delta < p_\pi$, then using sampling will turn cases that can settle into cases that cannot whenever $0.11 < q \leq 1$. Therefore, as long as the divergence in estimates is large enough, every such case will turn from feasible to impossible to settle when sampling is used.

b. *Effect on Settlement Amount*

The expected en masse settlement under Scenario I with the attorney in control is:

$$S^* = n[(p_\pi + p_\Delta)v + \alpha(C_D - C_p/r)]/2 \quad (9)$$

To determine the effect on the settlement amount, we must compare (9) with (6). It is easy to see that sampling increases the expected settlement amount if, as is very likely, $C_p/r > C_D$.

¹⁰⁹ If $q \leq \alpha(C_p/r + C_D)/v$, then the case can settle with or without sampling. If $q \geq (C_p/r + C_D)/v$, then the case cannot settle whether sampling is used or not.

B. SCENARIO II: SAMPLE PLAINTIFFS RECEIVE OWN VERDICTS AND COSTS ARE SHARED EQUALLY

1. When the Plaintiffs Control the Settlement Decision

a. *Effect on Settlement Feasibility*

When sample plaintiffs receive their own verdicts, a plaintiff's expected value of litigating through trial knowing that sampling will be used depends on whether the claim is high or low value. Since α is the probability a plaintiff will be selected for the sample and since a sample plaintiff receives her own verdict, w^H or w^L , and a non-sample plaintiff receives the sample average, v , the feasibility conditions with sampling become for high-value and low-value claims, respectively:

$$(p_\Delta - p_\pi)[\alpha w^H + (1-\alpha)v] + \alpha(C_P + C_D) \geq 0 \quad (10)$$

$$(p_\Delta - p_\pi)[\alpha w^L + (1-\alpha)v] + \alpha(C_P + C_D) \geq 0 \quad (11)$$

We must compare (10) with (1), and (11) with (2). Doing so and applying the same method as above yields the following results (where $q = p_\pi - p_\Delta$):¹¹⁰

- If $p_\Delta \geq p_\pi$, all high-value and low-value cases can settle, but the surplus is less with sampling for high-value claims. The surplus is less with sampling for low-value claims if $p_\Delta - p_\pi < (C_P + C_D)/(v - w^L)$ and greater with sampling if the inequality is reversed.
- If $p_\Delta < p_\pi$, then for all high-value cases and all α , a case that cannot settle without sampling also cannot settle with sampling. But there are cases where settlement is scuttled with sampling: a case can settle without sampling but not with sampling if $\alpha(C_P + C_D)/[\alpha w^H + (1-\alpha)v] < q \leq (C_P + C_D)/w^H$.¹¹¹

¹¹⁰ It is possible to derive parallel conditions that apply when information about w is asymmetric, just as in Scenario I. *See supra* notes 107-108.

¹¹¹ If $q \leq \alpha(C_P + C_D)/[\alpha w^H + (1-\alpha)v]$, then the case can settle with or without sampling. If $q > (C_P + C_D)/w^H$, then the case cannot settle whether sampling is used or not.

- If $p_\Delta < p_\pi$, then for all low-value cases and all α (and with $q = p_\pi - p_\Delta$), a case that cannot settle without sampling also cannot settle with sampling. But there are cases where settlement is scuttled with sampling: a case can settle without sampling but not with sampling if $\alpha(C_P + C_D)/[\alpha w^L + (1-\alpha)v] < q < (C_P + C_D)/w^L$.¹¹²

b. *Effect on Settlement Amount*

Under Scenario II, the expected settlement amounts with sampling become for high-value and low-value claims, respectively:

$$S^{H*} = \{(p_\pi + p_\Delta)[\alpha w^H + (1-\alpha)v] + \alpha(C_D - C_P)\}/2 \quad (12)$$

$$S^{L*} = \{(p_\pi + p_\Delta)[\alpha w^L + (1-\alpha)v] + \alpha(C_D - C_P)\}/2 \quad (13)$$

We must compare (12) with (3) and (13) with (4). It is clear from inspection that sampling reduces S^{H*} . Sampling increases S^{L*} if $p_\pi + p_\Delta > (C_D - C_P)/(v - w^L)$, which should usually be the case unless defendant's litigation costs greatly exceed the plaintiff's or the low-value case is very close to the population average.

2. When the Attorney Controls the Settlement Decision and Settles En-Masse

a. *Effect on Settlement Feasibility*

The feasibility condition with attorney control and sampling is:

$$(p_\Delta - p_\pi)v + \alpha(C_P/r + C_D) \geq 0 \quad (14)$$

This is the same as for Scenario I with the attorney controlling the settlement decision and settling en-masse. Therefore, the same results hold.

¹¹² If $q \leq \alpha(C_P + C_D)/[\alpha w^L + (1-\alpha)v]$, then the case can settle with or without sampling. If $q > (C_P + C_D)/w^L$, then the case cannot settle whether sampling is used or not.

b. *Effect on Settlement Amount*

The expected settlement without and with sampling are the same as for Scenario I, so the results are the same as well. Sampling increases the expected settlement amount if, as is very likely, $C_p/r > C_D$.

REEVALUATING COMPLEX MEDIATION GENERALIZATIONS

EDWARD BRUNET*

Several generalizations dominate the mediation discourse. When discussing mediation one often hears the almost mythic words of trust, confidentiality, expertise, and asymmetric information advantages held by risk neutral and data rich insurance companies. This short essay critiques these generalizations and exposes them as incomplete and erroneous.

Mediator expertise is elusive and not always necessary. Mediators frequently lack substantive expertise and exhibit only procedural expertise. Their expertise is only partial and may be minimal.

Confidentiality, often deemed central to a mediation, is similarly overblown. In truth, the mediators commitment to confidentiality is overstated. Most mediators act to filter and then redistribute important information gained in earlier caucus sessions. Such “noisy mediation” is central to mediation theory and indispensable to settlement. Mediator comments are often pregnant with new information hints. The stereotype that data rich insurers, repeat players in dispute resolutions possess an advantage in making and receiving offers is not universally true. The emergence of sophisticated and efficient networks of organized plaintiffs who operate to prevent insurers from controlling the mediation process undercuts this generalization.

I have mixed reactions to trust, often claimed a mediation essential. To be sure, trust remains a helpful and useful characteristic that plays a major role in settlement, particularly in the early stages of mediation. However, units of trust are difficult to create and do not guaranty a successful mediated settlement.

* Henry J. Casey Professor, Lewis and Clark. I thank Jessie Young for research help and Jeff Jones and Kate Lichter for comments. Any errors remain mine. Readers should be aware that I plead guilty to mediating over 75 disputes, mostly environmental insurance coverage and employment disputes.

I. INTRODUCTION

Several generalizations dominate mediation discourse, particularly in complex litigation. These commonly accepted generalizations, often involving insurance companies, have almost been transformed into myth and appear to be too widely accepted. In fact, these generalizations appear to lack reliable proof.

Consider the following generalized assertions regarding expertise, trust and success. For example, one often hears that the mediator is or should be an expert.¹ Expertise is the coinage with which we assess mediator hiring and competence. Similarly, the mediator appears clothed in a tunic of trust and is esteemed by the disputants because of such potential trust.² Trust has become a crucial ingredient of mediation and is the subject that takes center-stage in the parties' vetting process in mediator selection. Success represents yet another mediator homily. We often hear that a particular mediator is successful or, conversely, is no longer successful.³ Never mind that defining success which might be described in a variety of ways,⁴ including leading the disputants to dismiss a pending lawsuit but also just achieving new respect and a degree of self-awareness that will facilitate an ability to more properly evaluate settlement possibilities.

This list of mediation generalities goes on. One frequently hears that the presence of an insurance company skews the relative levels of information or, put more bluntly, creates information asymmetries.⁵ Some

¹ Leonard Riskin, *Understanding Mediators' Orientations, Strategies, and Techniques: A Grid for the Perplexed*, 1 HARV. NEGOT. L. REV. 7, 12 n.15 (1996) ("Most parties who are serious about resolving the dispute will choose a mediator who can give a strong, credible and objective evaluation of the legal and factual issues in the case.").

² See Kenneth R. Feinberg, *Mediation- A Preferred Method of Dispute Resolution*, 16 PEPP. L. REV. S5, S29 (1989).

³ See, e.g., Susan S. Silbey, *Mediation Mythology*, 9 NEG. J. 349 (1993); Riskin, *supra* note 1, at 12 n.15 (noting that Richard Ralston, a Kansas City lawyer-mediator, who has extensive experience as a lawyer and U.S. Magistrate judge, asserts that the disputants will select a mediator "who can close the negotiations.").

⁴ See Note, *Risk-Preference Asymmetries in Class Action Mediation*, 119 HARV. L. REV. 587, 608 (2005).

⁵ *Id.* at 595 n.39 ("[D]escribing how 'attorneys who run portfolios of cases (including class actions)' hope to 'diversify the risk' of new information causing bad outcomes in cases by assuming it will cause good outcomes in other cases in their portfolios.") (quoting Alon Harel & Alex Stein, *Auctioning for Loyalty*:

characterize insurance companies as data rich because they accumulate and save information from earlier disputes.⁶ As repeat players, insurers might have access to more information relevant to the dispute and be able to more accurately predict the probable outcome of a case.⁷ Of course, insurance companies invest in the business of collecting data and possess huge incentives to reap a return on this investment.

This advantage, or informational asymmetry, is allegedly of tremendous value when making settlement overtures to the possibly ill-informed plaintiff. In this world of complex mediation generalities, defendant insurers are often data-rich “repeat players”⁸ who know the score due often to access to much more information than their opponents regarding the past relevant judgments and settlements. Myth holds that these experienced insurers possess an important advantage in settlement negotiations because of their data-edge.⁹

Yet another generalization triggers the “Parable of Lucky Uncle Joe.” Plaintiffs or their attorneys have earned a contrasting generalized reputation for eschewing the data-based probabilities advanced in mediation or negotiation by defendants. Instead, the reigning generalization suggests that plaintiff or her attorney is prone to feel lucky and to ignore the data with the attitude that probability assessments are wholly wrong and “just won’t apply to me.” In this generalization the plaintiff Uncle Joe ignores the fact-based offer advanced by the data-rich, repeat player insurance company defendant because he feels lucky. Like the luck or hunch dominated fisherman who decides to go fishing with confidence on a day or a time that is unlikely to achieve success, Uncle Joe formulates his settlement offers and case analysis on hoped for good fortune rather than fact.

This essay exposes and critiques these mediation generalities. In a very real sense, these generalities seem to have derived from a type of

Selection and Monitoring of Class Counsel, 22 YALE L. & POL’Y REV. 69, 113 (2004)).

⁶ Gregory D. Ewig, *Using the Internet as a Resource for Alternative Dispute Resolution and Online Dispute Resolution*, 52 SYRACUSE L. REV. 1217, 1220-21 (2002).

⁷ *Id.* at 1223.

⁸ See Marc Galanter, *Why the “Haves” Come Out Ahead: Speculation on the Limits of Legal Change*, 9 L. & SOC’Y REV. 95, 97 (1974).

⁹ Blanca Fromm, *Bringing Settlement Out of the Shadows: Information About Settlement in an Age of Confidentiality*, 48 U.C.L.A. L. REV. 663, 701 n.173 (2001) (arguing that “repeat players in the settlement game” have the ability to absorb information about how particular types of cases settle).

mythology. Very little in the way of actual proof or data supports the generalities discussed. Instead, the propositions focused upon seem vague and can often be challenged as incomplete, questionable or overstated. I conclude by urging students and users of mediation services to avoid generality and, wherever possible, to challenge the use of hyperbole. While I am writing about mediation of complex cases generally, there is no question that the prime audience of this essay should be the insurance industry particularly. The generalities here addressed plague both counsel for defending insurers and their insurance clients.

II. GENERALIZATION #1: HIRE THE MEDIATOR WHO IS A SUBJECT MATTER EXPERT.

A great amount of ink has been spilled extolling the need for mediator expertise.¹⁰ Several problems exist with the notion that disputants should seek an expert mediator. First, determining the degree of mediator expertise is deceptively difficult. For disputants who have not used the particular mediator being vetted, the normal problem of lack of first hand experience routinely occurs. Unfortunately, it normally takes a considerable amount of time to assess the true level of mediator expertise and that time, of course, is spent in the mediation process itself, well after mediator expertise is assessed.

A related problem is that much of the information garnered regarding the possible mediator is entirely second-hand and indirect. This means that what a party learns about the mediator is not really based upon more reliable first-hand observation or direct assessment.¹¹ What one learns about expertise is all too frequently dependent on the filter of a third party who comes close to monopolizing accurate assessment.¹² Professor Robert Bone characterizes settlement process as a form of monopoly.¹³ This

¹⁰ See e.g., Riskin, *supra* note 1, at 46 (“The need for subject-matter expertise typically increases in direct proportion to the parties’ need for the mediator’s evaluations...[and] the kind of subject-matter expertise needed depends on the kind of evaluation or direction the parties seek.”).

¹¹ Fromm, *supra* note 9, at 698 (“The most thorough, useful, and exclusive resource of settlement is a person’s own firsthand knowledge of settlements.”).

¹² *Id.* (describing the “virtual monopoly over information about confidential settlements”).

¹³ See generally Robert G. Bone, *The Economics of Civil Procedure* 79-80 (2003) (asserting that “settlement is a bilateral monopoly”). The logic of this point seems obvious, but contracting parties are free to negotiate with other potential partners, provided that competition exists. Once two potential contract partners

problem pervades the process of mediator selection and is not unlike the all-to-common problem of building a case entirely upon circumstantial evidence.

Then there is the obvious difficulty of defining a measure of mediator expertise or competence. Substantive knowledge of the law underlying the dispute differs greatly from the procedural nuances that become good mediation practice. So-called “evaluative mediation” often stems from a mediator’s substantive knowledge.¹⁴ Some disputants might desire such “expertise,” while others may react negatively to anything beyond procedural sophistication.¹⁵

Questions of the mediator’s procedural expertise also should be addressed with a factually sensitive approach. Mediators develop specialties that can facilitate their roles. For example, some mediators are experienced in complex litigation and others bring a full plate of employment law to the ADR table. It seems obvious that disputants select the mediator who is an appropriate fit.¹⁶

Korobkin and Guthrie challenge the ability of lawyers to accurately value assets and question whether repetition of key negotiation dynamics will improve lawyer performance.¹⁷ Others fail to rely on the ability of the so-called repeat player to control valuable information

spend time and money assessing a possible deal, the sunk costs might rule out tuning away to another contractint entity.

¹⁴ See Leonard Riskin, *Decisionmaking in Mediation: The New Old Grid and The New New Grid System*, 79 NOTRE DAME L. REV. 1, 14 (2003); Leonard Riskin, *Mediator Orientations, Strategies and Techniques*, 12 ALTERNATIVES TO HIGH COST LITIG. 11 (1994).

¹⁵ See Jeffry Stempel, *The Inevitability of the Eclectic: Liberating ADR from Ideology*, 2000 J. DIS. RES. 247, 264 (2000) (“In practice, however, it appears that the most highly sought mediators are those who provide exactly this sort of evaluative feedback to the parties and use some measure of evaluation as part of their facilitation of reasonable party dialogue leading to settlement.”).

¹⁶ See, e.g., Maria R. Volpe, *Taking Stock: ADR Responses in Post-Disaster Situations*, 9 CARDOZO J. CONFLICT RESOL. 381, 389 (2008) (criticizing recent BP Gulf oil spill system because the infrastructure for conducting these large scale events was created and implemented by those mainly outside the field in the disaster areas).

¹⁷ Russell Korobkin & Chris Guthrie, *Psychology, Economics and Settlement: A New Look at the Role of the Lawyer*, 76 TEX. L. REV. 77, 83 (1997).

essential to a settlement.¹⁸ Taken together, these influential studies make assessments of mediator expertise elusive and difficult to achieve.¹⁹

III. GENERALIZATION #2: THE MEDIATOR WILL KEEP SECRET OR PRIVATE THE INFORMATION LEARNED DURING A CONFIDENTIAL CAUCUS CONTEXT.

Perhaps no other mediation generalization strikes such a sharp, focused note as the notion that the mediator will keep confidential all the information learned during the mediation session. This confidentiality pledge goes to the very core of a mediated dispute and is often reasserted during the mediation session by the mediator. The typical mediator will cover and promise the need for mediator confidentiality at the opening phase of the mediation and reaffirm his obligation to be confidential also in caucuses. These uses of confidentiality are internal to the mediation in that they recur within the mediation process at the opening and curing caucuses. They are much different uses of confidentiality than those “external” to the phases of mediation such as questions of mediation privilege that may arise following a mediation.²⁰

Belief in the sanctity of this platitude permits a party to freely discuss the true issues in the case and to disclose the “interest” of the disputant to the mediator.²¹ If such party disclosures fail to occur, the crucial information flow that fuels the mediation process atrophies and chances of settlement diminish. Aptly put by Ellen Deason, “if parties are to participate in mediation wholeheartedly, they need to have confidence that they can predict the extent to which their statements will be protected from disclosure.”²²

There is little doubt that many mediators scrupulously follow this respected generalization in what has been labeled an “understanding”

¹⁸ Fromm, *supra* note 9, at 698-700.

¹⁹ Stempel, *supra* note 15, at 265 (“If, in actual use of what is generally considered mediation, participants frequently prefer mediators who bring evaluative techniques to the process, it is needlessly bucking reality to insist that ‘real’ mediation must be devoid of any evaluative component.”).

²⁰ See generally Ellen E. Deason, *Enforcing Mediated Settlement Agreements: Contract Law Collides With Confidentiality*, 35 U. DAVIS L. REV. 38 (2001) (describing a post-mediation problem unlike the internal promises of confidentiality that occur during a mediation).

²¹ Ellen E. Deason, *Procedural Rules for Complementary Systems of Litigation and Mediation—Worldwide*, 80 NOTRE DAME L. REV. 553, 563-64 (2005).

²² *Id.* at 564.

model of mediation.²³ The model focuses on the mediator purely as a listener, and as a trainer whose goal is to empower the disputants to resolve their own conflicts. Yet, many mediators appear to depart substantially from this reluctance to “filter” and “leak” the forbidden and learned fruits of private caucusing. In a process labeled “noisy mediation,” these mediators often find themselves revealing some of the data learned in a prior caucus session.²⁴

In their 1994 classic article emphasizing the economic implications of mediation, Professors Brown and Ayres suggested that such a “noisy mediation” would significantly help to settle the dispute.²⁵ In contrast to the understanding-based model of mediation, an economic analysis emphasizes caucusing in analyzing the value a mediator brings to a negotiation by controlling the flow of information. Hidden information can be a major impediment to settlement. “By shuttling back and forth between meetings with individual disputants, mediators can collect and distribute private information.”²⁶ Brown and Ayres explain that mediators do this by sending “noisy translations of information disclosed during private caucuses.”²⁷ For example, a mediator might determine based on caucuses that there is a zone

²³ See, e.g., GARY FRIEDMAN & JACK HIMMELSTEIN, CHALLENGING CONFLICT: MEDIATION THROUGH UNDERSTANDING (2008) (asserting that the “understanding” model of mediation focuses on the parties as being in control of the mediation process and outcome; or the parties themselves are in the best position to find a solution because they are the ones who created, and are living in the problem context); Gary Friedman & Jack Himmelstein, *Resolving Conflict Together: the Understanding-Based Model of Mediation*, 4 J. AM. ARB. 225, 226-30 (2005) (asserting that increase of “understanding” regarding an adversary will help to resolve a dispute creatively).

²⁴ See, e.g., John W. Cooley, *Mediation Magic: Its Use and Abuse*, 29 LOY. U. CHI. L.J. 1, 69-70 (1997).

²⁵ See Jennifer Brown & Ian Ayres, *Economic Rationales for Mediation*, 80 VA. L. REV. 323, 329 (1994) (setting forth a theoretical need for the mediator to use leaked information that will be filtered and formulated by the mediator to aid in settling a case).

²⁶ *Id.* at 326.

²⁷ *Id.* at 328. See also Douglas E. Noll, *The Myth of the Mediator As Settlement Broker*, DISP. RESOL. J. 42, 46-47 (May-July 2009) (discussing the procedural and real-life implications of the fine line between trust and confidentiality and explaining how they must be broken to reach a settlement); Janis Sue Porter, *Mediation of Personal Injury Cases: Mediation Can Settle Most Personal Injury Cases*, 52 OR. ST. B. BULL. 34, 35 (Feb.-Mar. 1992) (describing the confidential nature of mediation and how it eventually must yield to be able to secure a settlement).

of agreement between the parties or that a set of trade-offs might bring the parties closer to agreement. “Revealing that there are gains from trade or that a particular set of trades might be acceptable to the other side has the effect of indirectly disclosing to each party some of the mediator’s private discussions with the other side.”²⁸ Without such filtering and information transfer the mediator will be the only person in the process to achieve true “understanding.”²⁹

This economic theory helps to explain the important aspect of mediation practice. While mediators need to keep certain information confidential, they also need to evaluate and then transfer new, valuable information to disputants. Some of the new “filtered” information transferred will not be in the same form as it was when the mediator gained access to it. As transferred to a disputant, selected information may be “noisy.” The mediator may take X, a piece of information learned in confidence from disputant A, and later ask disputant B “how would you react if your disputant A had decided to do X?” This process of filtering and conveying ideas to get negotiation movement and to transfer new information from a disputant lies at the heart of the mediation process.

Brown and Ayres observed that mediators acted as information brokers who collected valuable information about the strengths and weaknesses of a dispute and the relating settlement value. This theory of noisy mediation relies upon the mediator’s willingness to intentionally transfer data and to do so in a way that facilitates settlement. Strict confidentiality represents a lack of sharing of information between parties and “greatly decreases the likelihood that any claim will be filed.”³⁰

Some degree of noisy mediation appears essential to settlement. Nevertheless, some commentators criticize such behavior as “most problematic” because a mediator’s proposal will likely involve “a possible settlement option that implicitly contains messages about the preferences or

²⁸ Brown & Ayres, *supra* note 25, at 327. See also Fran L. Tetunic, *Mediation Myths and Urban Legends*, 82 FLA. B.J. 52, 52-53 (May 2008) (asserting that total confidentiality is a myth in Florida’s court-mandated mediation).

²⁹ See, e.g., Edward Brunet, Charles B. Craver & Ellen E. Deason, *ADR: THE ADVOCATE’S PERSPECTIVE* 231-37 (4th ed. 2011) (setting forth mediator goals of achieving fairness, respect and ability to understand one’s opponent as above all other mediation purposes).

³⁰ Laurie Kratky Dor, *Public Courts Versus Private Justice: It’s Time to Let Some Sun Shine in on Alternative Dispute Resolution*, 81 CHI.-KENT L. REV. 463, 488 (2006).

facts of the other party.”³¹ This criticism essentially accuses noisy mediators of unethical conduct, a serious accusation beyond the scope of this essay.

IV. GENERALIZATION #3: THE DISPUTANTS’ SELECTION AND FUNCTIONAL USE OF THE MEDIATOR IS BASED AND DEPENDENT ON TRUST

One frequently hears that trust is the foundation of mediation.³² Without trust, a mediation is doomed to failure. The presence of trust creates an open information environment in which the parties feel free to disclose their interests to the mediator. The building of trust provides a safe environment in which the mediator can increase information flow and make efficient use of trust.

All-star mediator Ken Feinberg maintains that neutrality and trust are essential characteristics of an effective mediator.³³ The mediation process can only work effectively when there is a trust relationship between the parties and their mediator, and where each party develops a quantum of trust allowing information to flow to the mediator.³⁴

These generalizations regarding trust are difficult to challenge. I salute and acknowledge the crucial role of trust in the mediation process. At the same time, however, the notion that trust can be massed produced in a cookbook, mechanical recipe that can be sold to all potential buyers strikes me as highly questionable. We are not all All-star mediators with a rich background and long resume capable of creating trust as early in the

³¹ Carrie Menkel-Meadow, *Ethics in Alternative Dispute Resolution: New Issues, No Answers from the Adversary Conception of Lawyers' Responsibilities*, 38 S. TEX. L. REV. 407, 443 (1997) (asserting that “[s]uch issues cannot be resolved easily either by broad protections of confidentiality or by reference to the lawyers' (and even other professionals') duties of confidentiality.”).

³² Feinberg, *supra* note 2, at S29 (Trust in a mediator is essential because if the parties believe that a mediator may be required to divulge information of the mediation, parties would be deterred from choosing mediation as a means of conflict resolution); *see also, e.g.*, Christopher Harper, *Mediator As Peacemaker: The Case for Activist Transformative-Narrative Mediation*, 2006 J. DISP. RESOL. 595, 602 (2006) (describing, *inter alia*, mediator trust and non-intervention or neutrality as a pervasive myth in mediation).

³³ Feinberg, *supra* note 2, at S29.

³⁴ *Id.* (“Any suspicion that the mediator may become an adversary or witness against one of the parties in future litigation will undermine the parties’ trust in the mediator.”).

mediator vetting stage. Respected mediators are few and far between and even they need to earn the nuggets of trust needed to grease the units of information flow essential to mediation success.

The timing associated with building trust merits careful attention. I stress that early trust building is essential to a winning mediation formula.³⁵ If high levels of trust are present early in a mediation context, the parties will offer the mediator critical information to be filtered and perhaps distributed in the form of noisy mediation comments by the mediator. However, the absence of any trust early in the mediation will delay information flow and related symmetrical data possession essential to resolving the dispute.

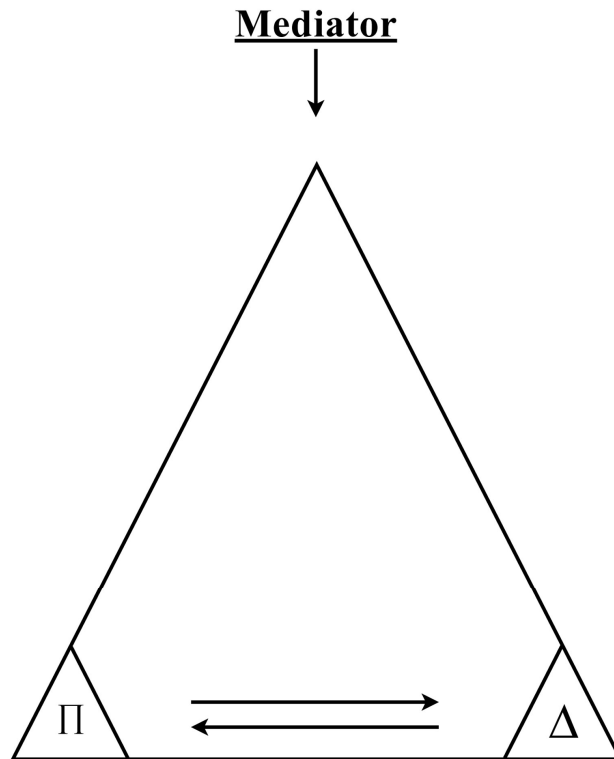
Trust is not created magically or mechanically. What factors cause the growth and emergence of trust essential to mediation? A brief review of scholarly thinking reveals a less than clear answer. Some mediators stress the value of reputation as the main path to enhanced trust.³⁶ In great contrast, others regard trust as an ill-defined foundational concept of mediation because “it hides the normative judgments that a mediator must make about what are good and bad agreements under the practical circumstances at hand.”³⁷

The relationship between trust and mediation can be depicted graphically. Consider the triangle below which illustrates the critical role of trust in the mediation process. The space within and without this triangle represents a negotiation context and is designed to help understand how the interjection of the mediator and units of trust into the disputing fray will heavily influence the settlement process.

³⁵ Christophe Leslie, *Trust, Distrust and Antitrust*, 82 TEX. L. REV. 515, 569 (2009) (noting that “goodwill gestures start the evolution of a relationship from suspicious competitors to trusting partners”).

³⁶ See, e.g., Feinberg, *supra* note 2; Leslie, *supra* note 35, at 570 (asserting that “honoring ones word and staying out of ones way” help trust building).

³⁷ See James R. Coben, *Gollum, Meet Smeagol: A Schizophrenic Ruminaton on Mediator Values Beyond Self-Determination and Neutrality*, 5 CARDOZO J. CONFLICT RESOL. 65, 73 (2004) (setting forth a less than enthusiastic view of trust in mediation).

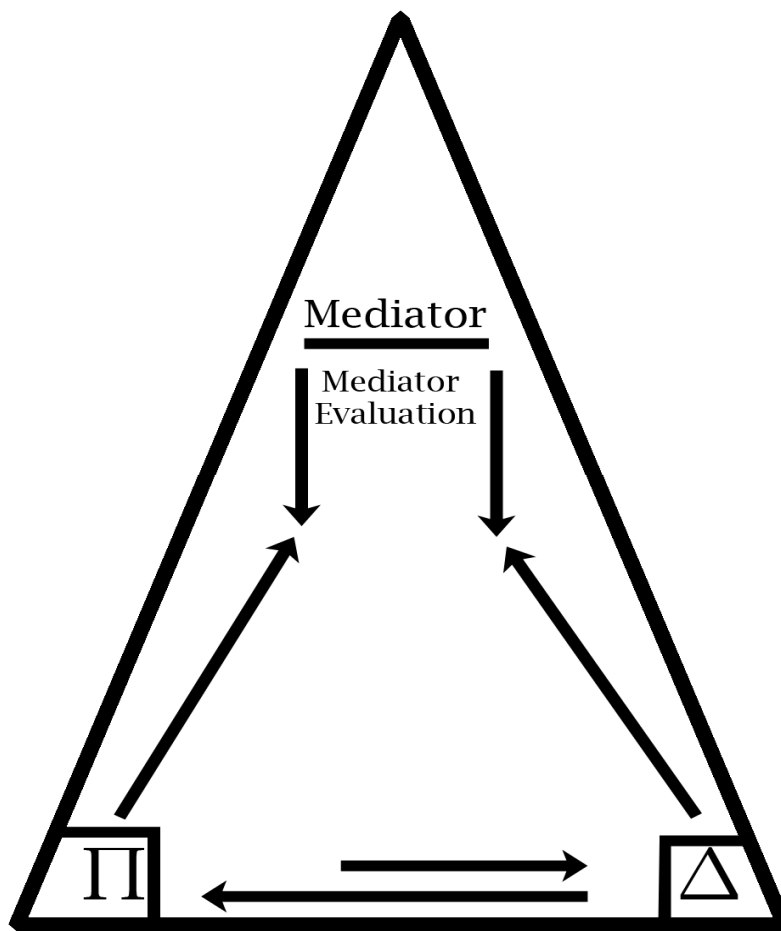
Graph #1: No Information, No Deal

In Graph #1 the parties are engaged in negotiation but unable to agree to a deal. The mediator is of little or no help because she is located far from the action and is playing “catch-up” due to a lack of information. The mediator is “out of it”, well outside the information-rich triangle and incapable of linking the adversaries without a better structural position and without additional intake of information and production of trust. In contrast, the parties in this graph are positioned inside the triangle, a position that yields access to incentives and information. They have more information than the mediator and the incentives to negotiate and ultimately resolve the dispute. Yet, they lack the high level of information essential to reduce risk and to agree to settle. No adequate information, no deal.

Graph #2, below shows an overall lack of trust. In Graph #2 below the parties do not trust the mediator nor one another. Not surprisingly, no trust effectively means no deal. Settlement fails to happen. Note the

structural position of the mediator who is now within the negotiation triangle where information might be traded and the disputed facts might become more clear. Some potential for settlement exists. Yet, the mediator, perhaps new to the process, has not created trust essential to settlement. No trust, no deal.

Graph #2: No Trust, No Deal

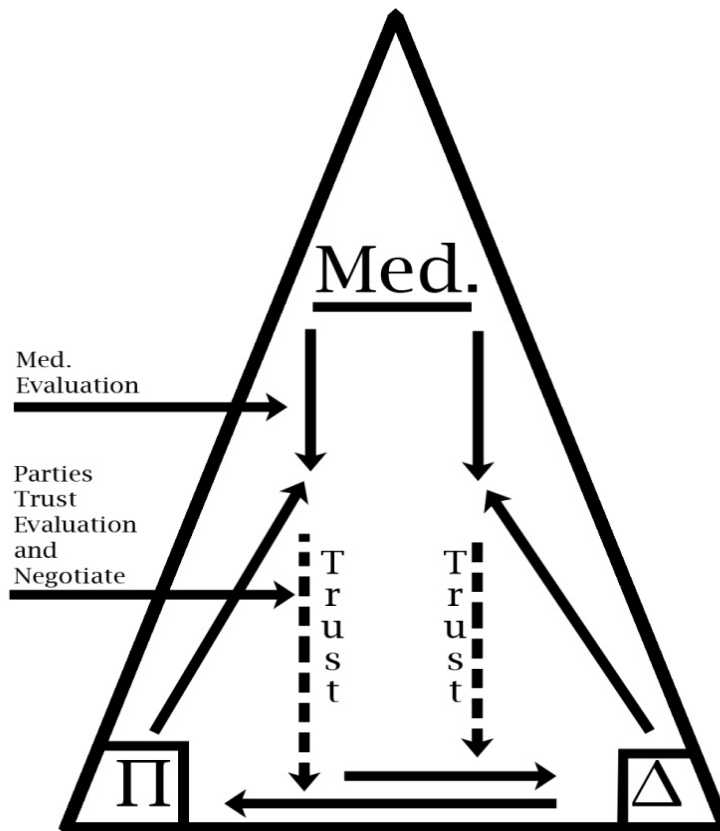


Graph #2, however, contains a structure that is supportive of settlement, which focuses on the need for both party respect and trust to make a deal. Of course, at times there will be the resources and a healthy

environment for agreement. In Graph #2 the mediator has assimilated additional information and is structured to aid in settlement inside the negotiation triangle. The mediator of Graph #2 is no longer “out of it” in terms of lacking information. Yet, a mediated deal remains impossible because trust seems nonexistent, preventing settlement.

In contrast, Graph #3, below, illustrates the impact of trust. The parties trust the mediator. Information appears to flow and risk may decrease.

Graph #3: Mediator Trust Results in Settlement



In Graph #3 the impact of mediation operates in a more positive way. The plaintiff and defendant are within the negotiation triangle, dealing

with each other. This dynamic is shown by the horizontal lines at the bottom of the triangle. Also note the position of the mediator within the triangle, meaning that the neutral is well-stocked with information and assessment abilities. The mediator is depicted as an insider who in Graph #3 is “with it,” able to help the parties deal with one another. The diagonal lines at the sides of the triangle show that the parties are playing the negotiation game with the mediator. This is inevitable. Yet, the parties are greatly aided by the presence of trust, here shown as a vertical line first connecting the adversary parties and then vertically influencing the negotiation itself by connecting the formerly distrustful adversaries.

The ingredient of trust seems the most important in the recipe of settlement. Peace is close at hand and trust has done its magic. The adversary parties trust and respect the mediator. In essence the graph demonstrates a mediator who enters into process well-informed (inside the negotiation triangle) and has the ability to produce trust which, in turn, facilitates settlement.

Here the adversary parties settle using probability assessments or values and the mediator is using evaluative mediation. The mediator hired is capable of collecting information to accurately assign probability assessments and appears well into the fray as an active participant. The mediator understands risk and is capable of communicating risks and evaluations to the disputants. Note also that the mediator has an incentive to help settle the case which should facilitate increases in mediator reputation.

The need to create trust is essential but elusive as well as complicated. Reputation appears the gold standard in the creation of early trust.³⁸ The achievement of a strong reputation as a neutral surely creates a degree of trust. The beauty of reputation-created trust is it can be triggered early in the multiple phases of litigation. Early trust based on reputation need not be prepared at or near the beginning of litigation. Reputation based upon trust exists without any need for a mediator to do much; it belongs to the mediator and typically is the mediator's to lose. Reputation will decrease where the parties observe untrustworthy actions.³⁹

There are other means of creating trust units, each difficult and each demanding. Some mediators strive for open and transparent events

³⁸ Ellen E. Deason, *The Need for Trust As A Justification for Confidentiality in Mediation: A Cross-Disciplinary Approach*, 54 U. KAN. L. REV. 1387, 1401 (2006) (“Reputation is one of the key variables in this calculus, for it will be enhanced by trustworthy behavior.”).

³⁹ *Id.*

held early in the dispute.⁴⁰ These events provide the disputants with an opportunity to see the mediator in action and hold the potential for either reducing or increasing trust. Caucusing will, of course, be undervalued in these events due to its placement of the mediator in secret positions that can worry the parties and decrease the quantum of mediator trust. Nonetheless, an early caucus gives a mediator opportunity to permit the clients to vent emotionally charged feelings and to begin to earn a reputation as an empathetic and active listener. These early trust enhancing events seem more appropriate for caucus contexts and are riskier if done in open session.

V. GENERALIZATION #4: INSURERS DOMINATE DATA RICH DISPUTES BECAUSE THEY POSSESS ASYMMETRIC INFORMATION AND LACK THE PLAINTIFF'S RISK-TAKER ATTITUDE

Data rich disputes, often defended by insurance companies, present the classic battle between parties differentiated by asymmetric information. One frequently hears the generalization that insurer's possess a rich mine of information which creates a significant advantage.⁴¹ The insurance industry files likely contain relevant information from previous disputes, such as settlement amounts of similar cases, judicial judgments, or mediator predispositions.⁴²

⁴⁰ See, e.g., Christopher Leslie, *Trust, Distrust, and Antitrust*, 82 TEX. L. REV. 515, 573-75 (2004) (noting that "Transparency Facilitates trust").

⁴¹ See Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1076 (1984) ("All plaintiffs want their damages immediately, but an indigent plaintiff may be exploited by a rich defendant because his need is so great that the defendant can force him to accept a sum that is less than the ordinary present value of the judgment . . . It might seem that settlement benefits the plaintiff by allowing him to avoid the costs of litigation, but this is not so. The defendant can anticipate the plaintiff's costs if the case were to be tried fully and decrease his offer by that amount. The indigent plaintiff is a victim of the costs of litigation even if he settles.").

⁴² See Howard M. Erichson, *The End of the Defendant Advantage in Tobacco Litigation*, 26 WM. & MARY ENVTL. L. & POL'Y REV. 123, 129 (2001) ("While in the past, one might have started with the assumption that the defendant had the resources to swamp the plaintiff, these [plaintiff] firms have accumulated sufficient capital through major victories in cases such as asbestos, tobacco, Dalkon Shield, etc., so that it may well be the plaintiff that is in the stronger resource position.").

Popular thought dictates that the insurer's asymmetric information creates a negotiation edge for defendants simply because they hold more relevant information. The myth posits that the insured defendant will be able to bargain more effectively by enabling more accurate assessments that are actually grounded in reality.⁴³

Part of this mythology suggests that the plaintiff and her attorney will take excessive risk and be classified as risk takers in decisions relating to trial or settlement. Consider the Parable of Uncle Joe, a risk-taking fisherman who tends to feel lucky and invincible. Uncle Joe is utterly confident and unable to compromise or to bargain realistically. He thinks that he will catch a big fish regardless of fishing conditions. Translating this Parable to the litigation context is not difficult. Some plaintiffs ignore bad news and turn down attractive offers by exercising risk-taking behavior.

But is the Parable of Uncle Joe universally accurate? Unlikely. Today's plaintiffs are represented by resourceful attorneys well connected with others with similar cases. Professors Issacharoff and Klonoff describe networks of plaintiffs' counsel who efficiently coordinate briefing and

procedural proficiency of mediation practice. Others selected as mediators may lack much substantive knowledge in the topic to be mediated. A healthy skepticism towards mediation expertise prevails and actually aids the vetting process.

Confidentiality, likewise, is a well-respected notion, almost a mantra of mediation practice, heard repeatedly at the beginning phases of a mediation. Like many clichés or mantras, this commitment to confidentiality seems to be either overstated or unrealistic. It fails to recognize that the mediator acts as a filter who analyzes and then distributes selected information learned in caucus session. Even Judge Posner has acknowledged the efficient use of noisy mediator comments that can facilitate settlement. The classic 1994 article by Brown and Ayres correctly justifies the process values of noisy mediator comments. A clear demand for evaluative mediation exists⁴⁵ and relies on this process of mediator assimilation and subsequent leaking or distribution of information collected from the adversary parties.

Moreover, the need for trust in the meditative process seems universally accepted and valid. It is hard to be critical of such a valuable and important tenet of dispute resolution. My comments on trust seem somewhat more measured and prudent. Mediation needs and depends on trust for its success.

Lastly, Generalization #4 combines two notions, one that the defending party, often represented by a data rich insurance company, possesses an advantage in making and receiving offers grounded in reality, and a second axiom, the Parable of Lucky Uncle Joe, the lucky fisherman-plaintiff who is a classic risk-taker who thinks that he can beat the odds habitually. When these two generalizations are combined, settlement mythology predicts that the defendant will evaluate settlement offers more accurately and will prepare and transmit to the mediator offers containing a patina of legitimacy. The mediator, in turn, may be impressed and thankful for this reality based information and might find herself subconsciously siding with the data rich party.

This scenario, while surely possible, occurs less with every passing day. Modern Uncle Joes appear, but with less frequency. Instead, individual plaintiffs' attorneys join forces to aggregate their discovery, monetary judgment information, and settlement data. Uncle Joe, while he still exists, is increasingly isolated. The nature of dispute resolution (or

⁴⁵ See Eric Green, *Reexamining Mediator and Judicial Roles in Large, Complex Litigation: Lessons From the Microsoft and Other Mega-Cases*, 86 B.U. L. Rev. 1171, 1186 (2006).

fishing) has changed greatly and now invites collaboration but on the plaintiff side of a dispute. Increased collective action has produced a new group of data-rich plaintiffs who are able to challenge the former information surplus held by insurance companies. Although I do not suggest that consortiums of collaborating plaintiff's attorneys hold information that is at the level of data-deep insurance companies, I do agree with those who have questioned the generalization that asymmetric information inevitably aids the insurance industry. Those times are past and a new era of collaboration on the plaintiff's side clearly has emerged.

PROBABILITY SAMPLING IN LITIGATION

JOSEPH B. KADANE*

Random sampling is a widely used and well-established techniques used to reduce the cost of providing interpretable data. This paper discusses examples in the several different kinds of litigation in which random sampling has been useful. The paper concludes with speculation about the possible use of random sampling in mass tort litigation.

This paper aims to contribute to a discussion of the possibility of using statistical methods to handle mass tort cases efficiently. After reviewing the basics of sampling, the paper summarizes cases involving sampling that the author participated in. The conclusion gives some thoughts on how mass tort litigation might be approached statistically.

I. PROBABILITY SAMPLING

The purpose of random sampling is to allow inference from the items observed to items unobserved. It is usually used to save the effort of having to observe each member of a population.

It is important to distinguish random sampling from other kinds of sampling. The hall-mark of random sampling is the use of a random number table or an equivalent computer program to choose units. The reason for the use of random numbers is to make transparent the process by which items are chosen for observation. This is important because without randomization, biases can creep in, whether advertent or inadvertent, that can destroy the validity of the inference to unobserved members of the population. While often random sampling is implemented in which each item has the same probability of selection, this is not necessary. What is necessary is that the probability of selection of each item be known in advance.¹

* Leonard J. Savage University Professor of Statistics and Social Sciences, Emeritus, at Carnegie Mellon University. His most recent books are "Statistics in the Law" (Oxford University Press 2008) and "Principles of Uncertainty" (Chapman and Hall 2011), also free on the web at www.stat.cmu.edu/~heidi/uncertainty-kadane.pdf.

Perhaps an example would illustrate this important point. Imagine a clinical trial of two treatments for a particular medical condition. Suppose that the physician (we'll call her Phyllis) who actually treats the patients observes the health of the patients in two categories, healthy and not, but this observation will not be available to those responsible for analyzing the results. Suppose also that healthy patients do better, whatever treatment is assigned to them, than do unhealthy patients. If Phyllis believes that one treatment is suitable for healthy patients and the other for unhealthy patients, and assigns treatments that way, the results of the trial will favor the treatment she assigns to healthy patients. If Phyllis wishes one treatment to be favored in the results, she can achieve this by her treatment assignments. In the first case her motives were pure, she was simply assigning treatments to help her patients as best she could. In the second case, her motives could be malign, for example, if she had a financial stake in her favored treatment. But her actions would be the same, and the consequences would be the same. Only by random sampling, where the decision of which treatment is assigned to a patient is removed from Phyllis, can outside observers be confident of lack of bias in the result.²

A relative of random sampling is systematic sampling, in which every k^{th} member of a list is used as a sample, starting with some arbitrary member of the list.³ Whether this is an adequate substitute for random sampling depends on the circumstances and the ordering of items in the list. Often the use of systematic sampling is benign. However, I remember one case in which systematic sampling was used to choose jury venires in Atlantic County, New Jersey.⁴ This has the effect that all persons with the same last name are adjacent in the list. Jurors were listed alphabetically starting with the fifth letter of their last name. There had been a previous system found to be discriminatory. A local bank proposed the following replacement. Often the choice of k was small, like 2 or 3. The consequence of this was that there were, more often than would have been true had the sampling been random, people in the same family, with the same last name, chosen for the same jury venire. Attorneys facing such a venire felt that they had to use peremptory challenges on every member of such a family if they challenged any member, to avoid offending potential jurors. The effect was

¹ See WILLIAM G. COCHRAN, SAMPLING TECHNIQUES 10-11 (J. Wiley & Sons 1977).

² For more on this example, see Scott M. Berry & Joseph B. Kadane, *Optimal Bayesian Randomization*, J. ROYAL STAT. SOC'Y B (1997).

³ See COCHRAN, *supra* note 1, at ch. 8.

⁴ State v. Long, 499 A.2d 264 (N.J. Super. Ct. Law Div. 1985).

to reduce the number of useable peremptory challenges available to the parties. This jury challenge was successful.

Sometimes there are special considerations that make it wise to separate a population of interest into various subpopulations, called strata, and to sample from each stratum. This, not surprisingly, is called stratified sampling. There are useful formulae to guide the choice of sample sizes for each such stratum.⁵

Another useful technique is sampling proportional to size. This is especially useful in sampling financial transactions in which the questions of interest center on dollar amounts rather than on typical items. Then if items are chosen for analysis according to the size of the transactions, a more accurate estimate of the dollar consequences of the transactions can result.⁶

Two standard references on random sampling are Cochran (1977)⁷ and Kish (1995).⁸

II. AN EARLY LEGAL EXAMPLE

Like new members of many organizations, new scientific methods go through a period of hazing by the legal system before accepted. For example, in *Sears, Roebuck and Co. v. City of Inglewood*, a random sample of days was selected to determine the proportion of sales made to non-residents of Inglewood (and therefore not subject to a sales tax). The best estimate from that sample was \$28,250 with a standard deviation of \$2,100, or a 95% confidence interval of \$24,000 to \$32,400 (per quarter for 11 quarters). The judge in the case rejected the sampling evidence, but permitted Sears to do a complete audit, which found the figure of \$26,750 per quarter (not counting some unavailable sales tickets).⁹

III. MORE RECENT EXAMPLES

This section is a brief survey of some cases that involve sampling, to display the wide variety of situations in which the technique is a cost-effective method of determining the approximate truth. I begin with

⁵ See COCHRAN, *supra* note 1, at ch. 5.

⁶ *Id.* at 250.

⁷ *Id.*

⁸ LESLIE KISH, *SURVEY SAMPLING* (J. Wiley & Sons 1995).

⁹ See R. Clay Sprowls, *The Admissibility of Sample Data into a Court of Law: a Case History*, 4 UCLA L. REV. 222 (1956-57) I did not participate in this case, and have no other source about it than Sprowls.

some disclaimers. First, this is nothing like a random sample of cases. Many of them are cases about which I have personal knowledge, because I was involved in them as an expert witness. Because many cases settle without a public record and few legal opinions say much about the sampling methods used, personal experience with cases seems an essential source of information. Second, some of the cases alluded to are still being litigated, and I am necessarily restricted in what I can say about them. Table 1 displays the topics to be discussed.

A. REMITTITURS

When a plaintiff has won a tort case, and damages have been awarded by a jury, the defendant can ask for a remittitur, under which the judge requires the plaintiff to accept a smaller damage award or a new trial, sometimes only on damages, sometimes on liability as well. The choice of a new trial seems required by the Seventh Amendment in federal cases, although this choice has been criticized as a sham.¹⁰ While the traditional criterion for awarding a remittitur is whether the jury award “shocks the conscience of the court”, New York, in a new law adopted in 1986 requires comparison with other similar cases. This requires the court to identify the cases it considers to be comparable, and then to analyze the amounts awarded to find the appropriate amount of remittitur in the case before it. Judge Weinstein, applying New York law, did this in the case of *Geressy v. Digital Equipment Corporation*.¹¹

There are several issues raised by this procedure. The first is the criteria used to determine comparability. A second is the database of cases available for study. This is usually cases of record, which omit cases that settled under conditions of confidentiality. Since plaintiffs are one-time players, while insurance companies are not, this asymmetry gives an incentive for secret settlement of cases with large damages. Third, when a list of comparable cases has been assembled, what remittitur should result?¹²

¹⁰ Suja A. Thomas, *Re-Examining the Constitutionality of Remittitur Under the Seventh Amendment*, 64 OHIO ST. L.J., 731 (2003); Joseph B. Kadane, *Calculating Remittiturs*, 8 L. PROBABILITY & RISK 125-31 (2009).

¹¹ *Geressy v. Digital Equip. Corp.*, 980 F. Supp. 640 (E.D.N.Y. 1997).

¹² On the latter point, see Kadane, *supra* note 10; Joseph B. Kadane, *Response to Professor Haug*, 8 L. PROBABILITY & RISK 137 (2009); Mark Haug, *Comment on Calculating Remittiturs by Kadane*, 8 L. PROBABILITY & RISK 133-35 (2009).

B. CROSS-SECTIONAL JURY CHALLENGES

A jury challenge is a motion to enforce the constitutional right to a jury venire composed of a representative cross-section of the community. Only some groups of people are considered “cognizable”, notably those based on race, sex and ethnic origin. Usually such a claim compares the proportion of a cognizable group in a series of venires to the proportion in the community often using census data.¹³ Data on the race, sex and ethnic origin of jury venires is often difficult to obtain, even concerning federal juries.¹⁴

C. STOPS ON THE NEW JERSEY TURNPIKE

The issue in this case is whether blacks were being stopped for traffic violations on the southern end of the New Jersey Turnpike at extraordinarily high rates.

A study from a stationary vantage-point (a bridge over the turnpike) yielded an estimate of 13.5% black drivers. A moving survey (from a car set on cruise-control at or near the speed limit) found roughly 15% black drivers, and that nearly all drivers were speeding, so the police, in principle, could stop whomever they wished. The proportion of black drivers among those stopped was about 46.2%, so the disparity was large, supporting a claim of differential enforcement of the law. The upshot was (1) evidence seized in about 15 stops was suppressed; (2) a consent decree with the Civil Rights Division of the Justice Department; and (3) some reform of the practices of the New Jersey State Police.¹⁵

¹³ For more on jury challenges generally, see David Kairys, Joseph B. Kadane & John P. Lehoczky, *Jury Representativeness: A Mandate for Multiple Source Lists*, 65 CALIF. L. REV. 776 (1977).

¹⁴ N. Chernoff & Joseph B. Kadane, *Preempting Jury Challenges: Strategies for Courts and Jury System Administrators*, JUST. SYSTEMS J. (forthcoming 2012).

¹⁵ See Joseph B. Kadane & Norma Terrin, *Missing Data in the Forensic Context*, 160 J. ROYAL. STAT. SOC’Y A 351-57 (1997); Joseph B. Kadane & John Lamberth, *Are Blacks Egregious Speeding Violators at Extraordinary Rates in New Jersey?*, 8 L. PROBABILITY & RISK 139 (2009); JOSEPH COLLUM, *THE BLACK DRAGON: RACIAL PROFILING EXPOSED* (2010); *State v. Soto*, 734 A.2d 350 (N.J. Super. Ct. Law Div. 1996).

D. WORKERS COMPENSATION INSURANCE

The law in many states requires employers to carry worker's compensation insurance, in case of an injury in the work-place. Private insurers offer such insurance, and participate in a high-risk pool in proportion to the premia for workers compensation insurance written by that insurer in that year. This gives each insurance company an incentive to under report. In a series of lawsuits, several insurance companies are accused of having done so, for example by attributing more premium to related auto and general liability insurance, so as to minimize their apparent workers compensation premium. Policy/years are being sampled to determine the truth of such allegations, and, if true, their extent. I serve as a court-appointed neutral expert to guide such sampling.

E. SALES TAXES

Pennsylvania sales tax excludes medication. Thus, Scope, which has no medication, is taxed, but Listerine, which has medication, is not taxed. The law requires retailers to collect sales tax. If the retailer fails to collect the tax owed, it must pay the missing tax to the state. If it erroneously collects tax, it must pay those funds to the state as well.

In a sales tax audit, the auditor told his team to be sure to include in the sample any Scope transactions they ran across. Thus, the sample was not random. In defense, I testified that I thought the retailer owed the \$6 found in uncollected tax, but not the \$300,000 the state wished to extrapolate from the \$6 they found. This case raises a general issue that the cost that might be gleaned from a random sample of transactions is a probability distribution for how much the taxpayer owes. But this does not specify how much the check should be.¹⁶

F. DISABILITY ACCESSIBILITY OF APARTMENTS

The Americans with Disabilities Act requires that apartments be accessible to the handicapped. To enforce this, architects sent to apartment complexes and select certain apartments to be assessed. If the selection of those apartments is not done by a random sample, the results cannot be reliably extrapolated to the apartments that were not inspected.

¹⁶ For commentary on this issue, see Joseph C. Bright, Joseph B. Kadane & Daniel S. Nagin, *Statistical Sampling in Tax Audits*, 13 LAW & SOC. INQUIRY 305 (1988).

G. INDIAN OIL AND GAS CLAIMS

The federal government has a fiduciary responsibility to collect royalties for oil and gas leases on Indian tribal lands. The tribes allege that it has not done so correctly, and have sued. To assess these claims, a random sample of leases is taken, and audited. An important difficulty is that the records kept by the Interior Department are incomplete.¹⁷

H. MEDICARE FRAUD

This case involved the defense of a physician who was accused of requiring medically unnecessary testing of patients in a laboratory he owned. The government wished to establish its case using a random sample of the patient records of the physician in question. Since the government's case was essentially an allegation of pattern or practice, it seemed that a random sample of carefully reviewed cases could be more informative than a hasty examination of every record. I was asked to testify that this sampling was an inherently unscientific approach, and the government should be required to examine every patient record. This I declined to do. It is possible to me that the law might require every patient record to be examined; scientifically a random sample of adequate size is sufficient.¹⁸ The defendant spent some time in prison.

IV. CONCLUSION

Random sampling is now widely used in litigation. Properly applied, it is an efficient way to find reasonable estimates of the facts, and the theory permits estimation of the sampling error.

There is interest in applying statistical sampling to mass tort litigation. In these cases, a large number of injured people are joined in a class, and liability has been found. The issue is how much to award to each person. Their circumstances and extent of injury (financial, physical, etc.) typically vary. The standard of the law, that each injured person deserves to have their individual case heard and judged, is administratively impossibly burdensome. Roughly the idea is to try a few cases, and use the outcome of

¹⁷ For one aspect of this work, see Mary S. Fowler & Joseph B. Kadane, *Oil and Gas on Indian Reservations: Statistical Methods Help to Establish Value for Royalty Purposes*, 14 J. STAT. EDUC. 3 (2006).

¹⁸ Joseph B. Kadane, *Ethical Issues in Being an Expert Witness*, 4 L. PROBABILITY & RISK 21 (2005).

those cases as guides to settle the rest. An argument is made¹⁹ that deliberate choice would better serve the ends of justice than the present system that allows the parties to speed or delay trials they deem to be helpful or harmful to their clients' interests. I believe that statistical ideas could be used in this setting, but just how to do it would depend on the specific context. I would look for variables that are believed to be important to determining the liability and the extent of damages. These might be used to create strata to be sampled from. More parsimoniously, a regression model (linear or non-linear) might be used. Until there is an actual case to address, these ideas should be taken as speculations.

Table 1: Brief description of the cases discussed

	Nature of Case	Legal Question	Sampled Items	Special Consideration	References
a.	remittiturs	is award in line with awards in comparable cases?	comparable cases	which cases are comparable? 7th amendment vs. due process	Thomas (2003); Kadane (2009a); Kadane (2009b)
b.	cross-sectional jury challenge	is the jury venire an adequate cross-section of the population?	jurors (race, sex, etc.)	date hard to get standards of adequacy	Jury work; Kairys, Kadane & Lehoczky (1977); NJ cases; Chernoff and Kadane (2011)
c.	stops on the NJ turnpike	racially differential law enforcement	drivers (race)	mind of officer missing race data	NJ v. Soto (1996); Kadane & Terrin (1997); Kadane & Lamberth (2009); Collum (2010)

¹⁹ Alexandra Lahav, *Rough Justice* (2011) (unpublished manuscript) (on file with the Connecticut Insurance Law Journal).

d.	workers compensation insurance	premiums appropriately reported to pool	WC insurance contract	availability of hard copy records	in litigation
e.	sales taxes	properly collected?	items sold and taxes collected	right of defendant to have all records examined	Bright, Kadane & Nagin (1988)
f.	disability accessibility of apartments	apartment complex in compliance with ADA?	apartment units	safe harbors	In litigation
g.	Indian oil and gas claims against federal government	proper collection of royalties	lease years	adequacy of records	Fowler & Kadane (2006)
h.	Medicare fraud	patient treatments appropriately billed	patient records	ethical issue	Kadane (2005)

IMPERMISSIBLE WINDFALLS?: UNEMPLOYMENT INSURANCE, BACK PAY, AND THE TWO CLASSES OF TITLE VII PLAINTIFFS

WYATT R. JANSEN*

I. INTRODUCTION

Prevailing plaintiffs in employment discrimination cases brought under Title VII of the Civil Rights Act of 1964 may be entitled to an award of back pay relief,¹ intended “to make the victims of discrimination whole by restoring them so far as possible . . . to the position where they would have been were it not for unlawful discrimination.”² Back pay relief under Title VII compensates plaintiff-employees for loss of pay attributable to discriminatory employment acts, including losses due to unemployment, underemployment, and failure to promote. Most common are suits alleging discriminatory firings,³ in which case back pay relief compensates, in whole or in part, for loss of income suffered during the period when the plaintiff was unemployed or underemployed due to an improper termination. Since many employees bringing Title VII firing suits qualify

* B.A., New York University, 2008; J.D. Candidate, The University of Connecticut School of Law, 2012. Many thanks to Professor Peter Siegelman.

¹ Civil Rights Act of 1964, Pub. L. No. 88-352, § 706(g), 78 Stat. 241, 261 (1964) (current version at 42 U.S.C. § 2000e-5(g)(1) (2006)); *see also* Albemarle Paper Co. v. Moody, 422 U.S. 405, 413-22 (1975) (holding that back pay, though an equitable remedy, may only be denied infrequently and for reasons that do not frustrate the purposes of Title VII).

² Ford Motor Co. v. EEOC, 458 U.S. 219, 230 (1982) (noting that this compensatory goal, while important, is a purpose that is secondary to the primary goal of Title VII to stop illegal employment discrimination) (citation omitted) (internal quotation marks omitted); *see* Griffith v. City of Des Moines, 387 F.3d 733, 746 (8th Cir. 2004) (“Title VII is not designed to provide a windfall to plaintiffs, but rather serves to put the plaintiff in the same position he or she would have been in absent discrimination.”).

³ *See* LAURA BETH NIELSEN ET AL., THE AMERICAN BAR FOUNDATION, *CONTESTING WORKPLACE DISCRIMINATION IN COURT 6* (2008), *available at* http://www.americanbarfoundation.org/uploads/cms/documents/nielsen_abf_edl_report_08_final.pdf (noting that 60% of all employment discrimination cases are brought because the plaintiff was fired, allegedly because of illegal discrimination).

for unemployment insurance,⁴ prevailing plaintiffs in such suits are likely to receive back pay awards that cover periods during which the plaintiff also received unemployment insurance benefits.

The overlap of back pay and unemployment insurance benefits described above is widely acknowledged to be a double recovery or “windfall” for plaintiffs.⁵ Consider the following illustration: employee “E”, living and working in Connecticut, suffers a discriminatory firing causing six weeks of unemployment. If E previously earned \$1,000 per week, her loss of pay from the firing is \$6,000, and she would likely collect \$6,000 in back pay under Title VII. If E also receives unemployment insurance benefits, she will be paid about \$462 for each week that she is unemployed,⁶ totaling \$2,772 over the six-week period of unemployment. Absent intervention, E collects a total of \$8,772 in compensation for the six-week period during which she actually lost \$6,000 of income. That is, from the perspective that unemployment insurance benefits stand in the shoes of a claimant’s ordinary wages, E actually lost \$3,288⁷ due to the discriminatory firing, and was overcompensated by the back pay award to the tune of \$2,722. On the other hand, if unemployment insurance benefits are not fully or truly paid for by employers, or if the benefits should not stand in the shoes of back pay as a matter of public policy, the \$8,772 in compensation may not be an overpayment.

⁴ A basic requirement to receive unemployment insurance benefits is that the applicant be involuntarily unemployed—a condition that an employee who is fired clearly meets. *See generally infra* Part II.B.

⁵ *See, e.g.,* Thomas W. Lee, *Deducting Unemployment Compensation and Ending Employment Discrimination: Continuing Conflict*, 43 EMORY L.J. 325, 335 (1994) (“[W]hile the deduction of unemployment compensation from back pay may provide a windfall for the employer . . . failure to offset unemployment benefits similarly provides a windfall for the employee.”) (footnote omitted) (internal quotation marks omitted).

⁶ *See* CONNECTICUT DEPARTMENT OF LABOR, ELIGIBILITY REQUIREMENTS – UNEMPLOYMENT INSURANCE, <http://www.ctdol.state.ct.us/progsupt/unemplt/uceligb.htm#Basic%20Eligibility%20Requirements> (last updated Oct. 11, 2011) (stating in Connecticut, a weekly unemployment insurance benefit entitlement is calculated by averaging the claimant’s income in the two highest of the four most recent quarters, and dividing that average by 26). Therefore, in E’s case, assuming a stable salary for the calculation period, E is entitled to $(((2 * \$12,000) / 2) / 26)$, or \$461.54 per week.

⁷ The difference between E’s ordinary weekly salary and her unemployment insurance entitlement.

This conflict has been divisive and remains unresolved by the Supreme Court. Absent controlling precedent, the lower federal courts have taken two distinct approaches.⁸ Some circuit courts of appeals have held that unemployment insurance benefits must be ignored when calculating a back pay award (the “restrictive rule”). In other circuits, the established rule allows district court judges to consider such benefits in their sound discretion (the “discretionary rule”), in which case the court may choose to either deduct unemployment insurance benefits from a back pay award or leave the back pay award undisturbed. (No circuit *requires* that the benefits be deducted.)

Much of the difference of opinion regarding the treatment of unemployment insurance benefits centers on whether or not those benefits are rightly considered “collateral sources.” Collateral sources are, in the simplest sense, benefits received by plaintiffs that are independent of—that is, collateral to—the defendant, and courts have traditionally been barred from considering such benefits when calculating a plaintiff’s damages.⁹ For example, a plaintiff who receives \$100 in support from his mother to compensate for a tortious loss of \$100 would be allowed under the collateral source rule to collect the full amount of damages from the tortfeasor, as those benefits were not sourced from, and are thus collateral to, the tortfeasor. Unemployment insurance benefits, on the other hand, are superficially *not* collateral to employers, since those employers are responsible for funding the unemployment insurance program.

Complicating this field further is what this Note terms “subrogation statutes,” which have been enacted in a significant minority of states. Subrogation statutes automatically reduce back pay awards by the amount of unemployment insurance benefits received during the same time period covered by a back pay award, and repay the recovered funds directly to the unemployment insurance fund.¹⁰ In those circuits with a discretionary rule, there is some evidence that district court judges consider whether or not a plaintiff will be subject to subrogation when calculating his or her back pay award.¹¹ On the other hand, district court judges in circuits following the restrictive rule are barred from considering the effect of subrogation.

⁸ See *infra* Part III.

⁹ See *infra* Part II.C.

¹⁰ See *infra* Part II.D.

¹¹ See *infra* Part III.B.2.

The effect of the circuit courts' approaches is illustrated in the following table,¹² assuming the same income and unemployment insurance benefit figures from the aforementioned illustration of plaintiff E:

Table 1 – Subrogation and the Circuit Courts				
	Plaintiff A	Plaintiff B	Plaintiff C	Plaintiff D
	Subject to subrogation, in a discretionary circuit.	Subject to a subrogation, in a restrictive circuit.	Not subject to subrogation, in a discretionary circuit.	Not subject to subrogation, in a restrictive circuit.
Total lost income.	\$6,000	\$6,000	\$6,000	\$6,000
Unemployment insurance benefits received.	\$2,772	\$2,772	\$2,772	\$2,772
Amount recovered by unemployment insurance fund through subrogation.	(\$2,772)	(\$2,772)	\$0	\$0
Back pay awarded by the court as damages in Title VII suit.	\$6,000	\$6,000	\$3,228	\$6,000
Total compensation received by plaintiff.	\$6,000	\$6,000	\$6,000 ¹³	\$8,772

¹² This table assumes that judges with the discretion to reduce a back pay award by unemployment benefits received will *always* reduce back pay in the absence of a subrogation statute and *never* do so when the plaintiff is subject to subrogation.

¹³ In the discretion of the district court judge, the \$6,000 back pay award is reduced by the amount of unemployment insurance benefits received, resulting in the prevailing plaintiff collecting a total of \$6,000 of both unemployment insurance benefits and back pay.

As is evident from the above table, three out of four combinations of circuit court approach and state law ensure that a plaintiff will receive the “right” amount of total compensation.¹⁴ But in states that do not have subrogation statutes in circuits following the restrictive rule, the prevailing plaintiff receives nearly 50% more compensation than otherwise similarly-situated plaintiffs.¹⁵ Though this result only occurs in one possible combination of state and federal law, a restrictive approach without subrogation is the governing legal standard in as many as twenty states, including California and Florida,¹⁶ and it is therefore likely that the majority of Title VII plaintiffs who have collected unemployment insurance benefits receive this double recovery.

This result is due exclusively to the complex and sometimes contradictory statutory and doctrinal frameworks that underlie this area of the law, particularly from the delegation of unemployment insurance regulation to the states and the resultant lack of a centralized policy regarding treatment of such benefits. This Note first discusses these discrete frameworks: Title VII,¹⁷ unemployment insurance, the collateral source rule, and state subrogation statutes. The approach by the federal appeals courts is subsequently discussed, as well as how the federal district courts exercise their discretion to consider back pay awards where they may lawfully do so. This Note then recommends an approach that may bring coherence to these inconsistent and often colliding structures and the approaches taken by the circuit courts, in the absence of a major reform of the unemployment insurance system.

II. THE STATUTORY AND DOCTRINAL STRUCTURE

¹⁴ That is, the back pay award that is necessary to replace the wages that the plaintiff actually lost due to a discriminatory employment action but not including unemployment insurance benefits, without regard for, as discussed *infra* Part II.B, the incidence of the unemployment insurance tax on employers.

¹⁵ This figure, of course, will vary based on factors including replacement rate, length of unemployment period, and salary. For example, since unemployment insurance benefits typically have an individual weekly benefit ceiling employees with high salaries will be overpaid by unemployment insurance benefits by much less than medium- and low-income plaintiffs as a proportion of their ordinary wages. See, e.g., CONNECTICUT DEPARTMENT OF LABOR, *supra* note 6 (in Connecticut as of October 2011, \$573).

¹⁶ See *infra* note 74, Part III.A.

¹⁷ For purposes of simplicity, this Note focuses on actions brought under Title VII, though the debate is relevant to other forms of employment discrimination, including suits arising under Section 1981 and the Americans with Disabilities Act.

This Note addresses the deductibility of unemployment insurance benefits from back pay awards in Title VII suits through analysis of the legal structures operating in the foreground and background of such cases. This Part will generally discuss the purpose of Title VII and remedies available under that statute, the system of unemployment insurance in the United States, the origins and rationale of the collateral source rule, and state subrogation statutes.

A. TITLE VII OF THE CIVIL RIGHTS ACT OF 1964

Title VII of the Civil Rights Act of 1964 prohibits discrimination on the basis of race, national origin, religion, or sex by private employers.¹⁸ Through express language, judicial interpretation, and congressional revision, the Act proscribes both intentional discrimination by employers as well as employment actions that lack discriminatory intent, but which have a disparate impact on persons from a protected class.¹⁹

While the Equal Employment Opportunity Commission is the official enforcement agency for Title VII,²⁰ the Commission brings only a small fraction of the employment discrimination cases that it reviews. It follows that most Title VII suits are brought by individuals hiring private counsel or proceeding *pro se*.²¹ As a result, plaintiffs have an important role under Title VII as private attorneys general, both asserting their individual right to be free from discriminatory employment actions and policing employers to vindicate the broader purposes of the statute—namely, to eliminate employment discrimination.²²

¹⁸ Civil Rights Act of 1964, Pub. L. No. 88-352, § 703, 78 Stat. 241, 255 (1964) (current version at 42 U.S.C. § 2000e-2 (2006)); *Griggs v. Duke Power Co.* 401 U.S. 424, 429-30 (1971); GEORGE A. RUTHERGLEN, *EMPLOYMENT DISCRIMINATION LAW* 6 (2d ed. 2007).

¹⁹ See generally RUTHERGLEN, *supra* note 18.

²⁰ See *id.* at 8.

²¹ See NIELSEN, *supra* note 3, at 15 (noting that the EEOC intervenes as plaintiff in just 3% of all employment discrimination cases).

²² See Donald T. Kramer, *Factors or Conditions Said to Justify Increase in Attorney's Fees Awarded Under § 706(k) of the Civil Rights Act of 1964*, 140 A.L.R. FED. 301 (1997) (stating the private attorney general model serves as the justification for Title VII's fee shifting structure, which awards attorney's fees to prevailing plaintiffs, but not to prevailing plaintiff's under ordinary circumstances).

Section 706(g) of Title VII permits courts to award back pay as an equitable remedy for illegal employment discrimination.²³ Though Title VII back pay sounds in equity and the plain language of Title VII is permissive, the Supreme Court has indicated that judges are significantly limited in their discretion to decide *not* to award back pay relief.²⁴ In addition to back pay, prevailing plaintiffs in Title VII suits have available a broad range of statutory relief: reinstatement or, if reinstatement is impossible, front pay;²⁵ additional compensatory damages for both pecuniary and non-pecuniary losses like job search expenses, reputational damage, and emotional pain and suffering;²⁶ punitive damages in circumstances where a defendant acts with malice or reckless indifference to the federally-protected rights of the plaintiff;²⁷ and finally, reasonable attorney's fees.²⁸

Part and parcel of using back pay as the primary remedy under Title VII is that suits brought under the statute tend to be low in value.²⁹

²³ Civil Rights Act of 1964, Pub. L. No. 88-352, § 706(g), 78 Stat. 241, 261 (current version at 42 U.S.C. 2000e-5(g)(1) (2006)) (“If the court finds that the respondents has intentionally engaged in or is intentionally engaging in an illegal employment practice, the court may . . . order such affirmative action as may be appropriate, which may include, but not limited to, reinstatement, or hiring or employees, with or without back pay”); *see Albemarle Paper Co. v. Moody*, 422 U.S. 405, 415-22 (1975) (stating back pay has become the presumptive remedy for employment discrimination).

²⁴ *See Albemarle Paper Co.*, 422 U.S. at 421 (“[B]ackpay should be denied only for reasons which, if applied generally, would not frustrate the central statutory purposes of eradicating employment discrimination throughout the economy and making persons whole for injuries suffered through past discrimination.”).

²⁵ 42 U.S.C. § 706(g), 78 Stat. at 261 (current version at 42 U.S.C. § 2000e (2006)).

²⁶ The Civil Rights Act of 1991, Pub. L. No. 102-66, § 102(b), 105 Stat. 1071, 1073 (1991) (current version at 42 U.S. 1981a (2006)).

²⁷ *Id.* (stating punitive damages are only available in certain forms of employment discrimination cases and subject to the same caps as compensatory damages); *see Kolstad v. American Dental Ass’n*, 527 U.S. 526, 529-40 (1999) (stating the standard for whether punitive damages are appropriate is not egregiousness, but rather whether the employer has engaged in discriminatory act despite perceiving that the act is in violation of federal law).

²⁸ The Civil Rights Act of 1991, Pub. L. No. 102-166, § 103, 105 Stat. 1071 (1991) (codified as amended at 42 U.S.C. § 1988(b)).

²⁹ *Discrimination Law in the 1990’s*, in HANDBOOK OF EMPLOYMENT DISCRIMINATION RESEARCH: RIGHTS AND REALITIES 261, 265 (Laura Beth Nielsen

While the size of damages awards increased after the 1991 amendment of Title VII, which made available compensatory relief, punitive damages, and jury trials,³⁰ the median Title VII back pay award remains under \$50,000.³¹

B. UNEMPLOYMENT INSURANCE IN THE UNITED STATES

The Social Security Act of 1935 first enabled the unemployment insurance system. Rather than creating a federal regime, the Act instead encouraged the states to enact their own unemployment insurance programs, so long as they operated within certain federal guidelines (such as minimum tax rates).³² This joint federal and state statutory scheme has resulted in state unemployment insurance programs that are often widely divergent with respect to coverage, benefits, funding, and administration.³³

Despite this divergence, there are points of congruence among the state systems. Across all states, unemployment insurance benefits share a common aim of providing partial and temporary wage replacement for involuntarily unemployed workers meeting certain conditions regarding continuity and type of employment.³⁴ These benefits primarily serve two goals: narrowly, to stabilize the standard of living for unemployed individuals during those individuals' periods of unemployment, and broadly, to reduce overall economic volatility during periods of widespread unemployment.³⁵

& Robert L. Nelson eds., 2005) (showing that pre-1991, employment discrimination back pay awards were usually small and positively related to an employee's pay).

³⁰ *Id.* at 268 tbl.3.

³¹ *Id.* at 279 tbl.8. The \$50,000 figure is one that is not per plaintiff, but rather per case; because a not-insignificant number of cases are brought with more than one plaintiff. *See* NIELSEN, *supra* note 3, at 12 (6% of employment discrimination cases involved 2-10 plaintiffs); it is likely that plaintiffs' actual awards are somewhat lower.

³² Edwin E. Witte, *Development of Unemployment Compensation*, 55 YALE L.J. 21, 22 (1945); U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-10-440, UNEMPLOYMENT INSURANCE TRUST FUNDS 3 (2010).

³³ Gillian Lester, *Unemployment Insurance and Wealth Redistribution*, 49 UCLA L. REV. 335, 344 (2001).

³⁴ *See id.* at 344-46.

³⁵ *Id.* at 341-42.

The unemployment insurance system is directly financed by both state and federal payroll taxes, paid solely by qualifying employers,³⁶ which together constitute as much as 15% of an employer's total annual tax bill.³⁷ The structure of tax rates is twofold: first, a federal rate that is as low as .8% and is applied to a base of an employee's first \$7,000 of wages; second, a state tax, with rate and base terms that vary widely among the states, but that must remain within certain federal guidelines.³⁸ These state tax rates are typically adjusted annually and calculated relative to a state's unemployment insurance fund balance, with a lower balance triggering higher overall rates and a higher balance resulting in generally lower rates.³⁹

Through a process known as the "experience rating," market-wide state tax rates are adjusted for each employer based on that employer's history of firing its employees, with the resultant individualized tax rate

³⁶ With the notable exceptions of Alaska, New Jersey, and Pennsylvania, which withhold unemployment insurance taxes from employee wages in addition to taxing employers. See U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 32, at 4 n.9. For the purposes of simplicity and coherence, this Note ignores these exceptions to the general rule.

³⁷ Lester, *supra* note 33, at 340. The employer-funded model used by the United States is distinctly different from the financing of unemployment insurance in other countries, where funds come from a variety of sources exogenous to employers. See Steven Jurajda, *Unemployment Outflow and Unemployment Insurance Taxes*, CERGE-IE Working Paper Series No. 143 at 2 (1999).

³⁸ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 32, at 5.

In Connecticut, for example, the tax rate for newly established employers is 3.7% as applied to the first \$15,000 of each employee's wages, with rates for established employers ranging from 1.9% to 6.8%. See Employer Information Notice, CT Unemployment Insurance Tax, Connecticut Department of Labor, Sept. 2011, available at <http://www.ctdol.state.ct.us/uitax/EmplNotices/EmplNotice0911.pdf>. Thus, newly-established Connecticut employers employing employees making \$15,000 or more per year pay \$555 in unemployment insurance taxes per employee per year. See *id.*

In Texas, for another example, the rate for a new employer is .78%, with maximum and average tax rates of 8.25% and 2.03%, respectively, applied to a base of \$9,000 of wages. Unemployment Tax Rates, Texas Workforce Commission, <http://www.twc.state.tx.us/ui/tax/unemployment-tax-rates.html>. Thus, an employer would pay (per year and per employee earning \$9,000 or more in annual wages) unemployment insurance taxes of \$70.2 at the minimum rate and \$182.70 at the average rate.

³⁹ See U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 32, at 7.

called the “experience-rated component.”⁴⁰ Through use of the experience rating, an employer with a history of many firings (and one that has thus imposed a high cost on the unemployment insurance pool) is subject to a higher tax rate than an employer without such a history. Use of the experience rating can adjust the effective tax rate on an employer to as high as 10.5%,⁴¹ though states typically set a maximum rate. Because the experience rating can thereby impose significant financial consequences on an employer for firing employees, the rating is believed to have the effect of deterring layoffs,⁴² and may be a means of controlling employer-side moral hazard in unemployment insurance generally.

To be sure, the cost of unemployment insurance benefits may be indirectly paid, in whole or part, by employees in the form of reduced earnings—that is, the ultimate incidence of the unemployment insurance tax may fall on employees. The significance of tax incidence is that, to the extent that the cost of unemployment insurance is borne by employees instead of employers, the benefits are less clearly categorized as independent of (that is, *not* collateral to) employers, and public policy and collateral source rule doctrine thus may more strongly favor treating the benefits as collateral.

The unemployment insurance system appears designed for the incidence of the tax to apply fully to employers, by not requiring contribution from employees and through use of the experience rating.⁴³ Despite this intention, however, it may be that employers shift the incidence of the unemployment insurance tax forward, by charging consumers more for goods or services, or backwards, by reducing the price they pay for labor input.⁴⁴ There is no consensus that the unemployment insurance tax is back-shifted, and those studies that have attempted to isolate the effect of the unemployment insurance tax have come up with divergent results.⁴⁵ Recent data suggests that the costs may be shared between employers and employees, with one study showing that employers

⁴⁰ *Id.* at 5.

⁴¹ Lester, *supra* note 33, at 345 (in Pennsylvania).

⁴² See Jurajda, *supra* note 37, at 2 (the experience-rated component has been demonstrated to influence employer decision-making in regards to both initially laying off workers and recalling previously laid-off workers).

⁴³ See generally Lester, *supra* note 33.

⁴⁴ *Id.*

⁴⁵ See *id.* at 382.

are not able to shift the costs of inter-firm experience rating variances, but may be able to shift some portion of the “market rate”, or base tax burden.⁴⁶

Cost shifting, if it does occur, may not be so simple in the case of unemployment insurance, however. Since the experience rating means that a firing costs an employer money in the form of higher tax rates, and, as discussed above, that upward adjustment is not shown to be back-shifted, any employee contribution to the unemployment insurance tax may be negated by the cost of his firing to the employer. For example, Texas calculates its experience rating by dividing the last three years of unemployment insurance benefits paid out over three years of an employer’s taxable wages, and multiplying that by a flat tax rate.⁴⁷ For example, suppose a new Texas employer employs three workers at \$10,000 per year for a period of three years, but fires one worker at the end of year 2, entitling that worker to collect a 50% unemployment insurance benefit. That employer’s effective unemployment tax rate will resultantly increase from 2.72% to 8%;⁴⁸ on the Texas taxable wage base of \$9,000, the employer would pay \$1,440 in unemployment insurance tax in year three per employee, as compared to \$482 in year two. That is, the firing will cost the employer nearly \$1,000 per employee per year for the three years that the firing is computed in the employer’s experience rating. This is all to demonstrate that the back-shifting of tax may be mitigated by the

⁴⁶ Patricia M. Anderson & Bruce D. Meyer, *Effects of the Unemployment Insurance Payroll Tax on Wages, Employment, Claims and Denials*, 78 J. PUB. ECONOMICS 81, 95 (2000) (noting, however, that “large standard errors preclude [the authors] from drawing strong conclusions”).

In Texas, for example, the “market” tax rate is .78%, while the average experience rating tax rate is 1.96%, both applied to \$9,000 of wages. If employees are responsible for the entire market rate (as opposed to partial responsibility, as demonstrated by the above-referenced study), and employers were responsible for the experience rate, the proportion of employee to employer contribution would be roughly 1:3. See Texas Workforce Commission, *supra* note 38.

There is an intuitive logic to the findings by Anderson and Boyer, if back-shifting does in fact occur. Since the experience rating is determined based on an employer’s past history of layoffs, firms that reduced employee wages to account for higher unemployment taxes that result from the experience rating would essentially expect employees to be paid less to work for an employer that is more likely to cause them to be unemployed.

⁴⁷ See TWC TAX DEPT., *Your Tax Rates – 2011*, <http://www.twc.state.tx.us/ui/tax/unemptax2011.html#oa> (Sept. 30, 2011).

⁴⁸ *Id.* (That is, [$\$5,000/\8000] * [1.28]).

concomitant increase in tax burden caused by a firing, which itself appears to be fully absorbed by the employer.

Unemployment insurance funds have been in financial peril in recent years. At the close of the fourth quarter of 2009, state unemployment insurance fund balances were the lowest they have ever been in the history of unemployment insurance,⁴⁹ and this undercapitalization is expected to worsen with the ongoing recession.⁵⁰ Compounding these historically low funding levels is the reality that loans from the federal government are currently buoying the balances of many state unemployment funds; because these loans are reflected in the historically low fund balances, state funds are likely even more weakly positioned than they appear at first glance.⁵¹

C. THE COLLATERAL SOURCE RULE

Circuits that follow the restrictive rule typically do so on the basis that considering unemployment insurance benefits violates the collateral source rule. As a general principle of tort law (treated as both a rule of evidence and as substantive law), the collateral source rule proscribes courts from considering benefits received by a plaintiff that are independent of (i.e., collateral to) a defendant when calculating a plaintiff's damages.⁵² The collateral source rule typically does not protect benefits provided by a defendant or a party identified with a defendant,⁵³ but rather only applies to truly independent or third-party sources, such as gratuitous support from family members or an unintended benefit arising from a defendant's wrongful act.⁵⁴

Perhaps the quintessential application of the collateral source rule is to exclude evidence of plaintiff-purchased insurance benefits covering a loss for which that plaintiff is later awarded damages. For instance, due to a negligent act by tortfeasor "T", plaintiff "P" incurs medical expenses that are covered by P's insurer. In a later suit against T, P may still collect damages for medical expenses when the collateral source rule is applied, even though P did not pay for those expenses out of pocket. In this

⁴⁹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 32, at 9 (after figures were adjusted for inflation).

⁵⁰ *Id.* at 13-14.

⁵¹ *See id.* at 9 (overall balance of state fund reserves was negative \$15 billion).

⁵² BLACK'S LAW DICTIONARY 299 (9th ed. 2009) (defining Collateral-Source Rule).

⁵³ Linda L. House, *Section 1983 and the Collateral Source Rule*, 40 CLEV. ST. L. REV. 101, 103 (1992).

⁵⁴ STEVEN SHAVELL, *ECONOMIC ANALYSIS OF ACCIDENT LAW* 143 (1987).

circumstance, P effectively receives a double recovery—one from the insurer and another from the tortfeasor, both purporting to compensate for the same injury. It is in reaction to such results that proponents of “tort reform” efforts have sought to abrogate the collateral source rule.⁵⁵

This common application of the collateral source rule may not in fact result in a windfall,⁵⁶ however, despite the mechanism of P receiving two payments for one injury. In the above example, P purchased insurance coverage and paid insurance premiums in exchange for the contractual right for payment upon the occurrence of the tortious act at issue. Because the insurance coverage was purchased in order to cover the cost of the injury, presumably the premiums were calculated so as to pay for the cost of covered events, plus overhead.⁵⁷ By charging P a rate calculated to the risk and cost of a covered event, the insurance policy primarily changes the timing of the payment for the injury to the period when P makes premium payments, but does not alter that P has a cost associated with the injury that will require compensation in order for P to be made whole.

Further preventing a windfall in many traditional insurance applications of the collateral source rule is the effect of a subrogation right, held by many or even most insurers.⁵⁸ Where this right exists, a collateral source is entitled to the rights and remedies belonging to the plaintiff for which the plaintiff was compensated by the collateral source, eliminating any windfall *ex post*.⁵⁹ Thus, if P’s insurance contract includes a subrogation right (and it likely does), his insurer may seek to collect the medical damages awarded to P to the extent that it reimbursed P for such

⁵⁵ See Jamie L. Wershbale, *Tort Reform in America: Abrogating the Collateral Source Rule Across the States*, 75 DEF. COUNS. J. 346, 349 (2008).

⁵⁶ For the purposes of this note, “windfall” and “double recovery” are distinguished. Windfall will describe the situation where a plaintiff collects more than his actual losses, and is thus overcompensated or “profits” from a defendant’s wrongdoing. Double recovery will mean that a plaintiff receives two payments for the same injury, which, as this Note explores, may or may not result in a windfall. See BLACKS LAW DICTIONARY, 1738 (9th ed. 2009) (defining windfall as “[a]n unanticipated benefit, usu. in the form of a profit and not caused by the recipient”). Compare *id.* at 1389, *Double Recovery* (“a judgment that erroneously awards damages twice for the same loss . . . [or] recovery by a part of more than the maximum recoverable loss . . .”).

⁵⁷ RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 253 (8th ed. 2011).

⁵⁸ Wershbale, *supra* note 55, at 349-50 (noting, however, that subrogation rights are rarely asserted).

⁵⁹ See *id.* at 349.

expenses, and P will (theoretically) have paid lower premiums to account for this.⁶⁰

Aside from factors that mitigate the occurrence of a windfall, there are public policy justifications that favor the collateral source rule even when it does result in a windfall. Windfalls can serve as a rough means of providing for attorney's fees, which, if not otherwise available, would detract from the make-whole nature of compensatory relief.⁶¹ Similarly, a windfall may be used to award punitive damages when they are not provided by law.⁶²

Perhaps more convincing are those policy arguments related to the redistributive and deterrent uses of windfalls. For one, allowing insurance coverage to reduce a wrongdoer's cost for his wrongs reduces the concomitant incentive to prevent future wrongdoing to avoid future costs of similar wrongs.⁶³ The collateral source rule thus furthers the deterrent function of compensatory relief.⁶⁴ There is also an intuitive preference to award windfalls, where they must exist, to the victim and not the violator.⁶⁵ Finally, if insurance reduces a plaintiff's tort award dollar-for-dollar, there is significantly less reason to buy insurance in the first place, and there are strong reasons for favoring insurance coverage.

There are, however, several competing considerations. Any windfall may be inappropriate in a make-whole relief scheme, which is focused on compensating plaintiffs for actual losses, and is less concerned with the source of that compensation.⁶⁶ Where statute or other relevant law speaks clearly on the issue, using the collateral source rule to roughly

⁶⁰ POSNER, *supra* note 57, at 253.

⁶¹ That is, a plaintiff with paid counsel will either pay an hourly fee or will have a contingency agreement, costs of which may not be accounted for in an ordinary damages award that is not accompanied by attorney's fees. Since unreimbursed fees either indirectly (in the case of an hourly rate) or directly (in the case of a contingency agreement) reduce the amount of damages actually recovered by the plaintiff, the award may no longer put the plaintiff in the position he or she would have been in but for the wrongdoing. *See, e.g.,* Daena A. Goldsmith, *A Survey of the Collateral Source Rule: The Effects of Tort Reform and Impact on Multistate Litigation*, 53 J. AIR L. & COMM. 799, 802 (1988); Robert Hernquist & Arthur v. Catour, *An Examination of the Collateral Source Rule in Illinois*, 38 LOYOLA U. CHI. L. J. 169, 172-73 (2006).

⁶² *See* Hernquist, *supra* note 61, at 181.

⁶³ POSNER, *supra* note 57, at 253.

⁶⁴ Goldsmith, *supra* note 61, at 801; POSNER, *supra* note 57, at 253.

⁶⁵ House, *supra* note 53, at 104; Hernquist, *supra* note 61, at 188.

⁶⁶ Hernquist, *supra* note 61, at 182.

provide attorney's fees and punitive damages may be unnecessary or improper.⁶⁷ Finally, subrogation rights may be an inefficient solution to preventing windfalls,⁶⁸ and the collateral source rule may, by increasing damage awards, artificially inflate insurance premiums.⁶⁹

Perhaps as a result of these concerns, the collateral source rule has been steadily weakened: erosion of the rule began almost immediately following its adoption by the Supreme Court in the 19th Century,⁷⁰ and by 2007 all but 12 states had created some statutory alteration to the common law rule.⁷¹ For example, a number of states now allow post-verdict reduction of a defendant's liability for collateral benefits received that are not subject to subrogation, while refusing such a reduction where the collateral source does hold a subrogation right.⁷²

D. STATE SUBROGATION STATUTES

States have taken several statutory approaches to the double recovery that can result when back pay awards overlap with unemployment insurance benefits,⁷³ but by far the most common is to invest a legal subrogation right in the state's unemployment insurance fund. At least 16 states accomplish this through statutes requiring reimbursement of state unemployment funds for insurance benefits paid that overlap with back pay awards, typically by the employer repaying the fund directly and then

⁶⁷ See *id.* at 186.

⁶⁸ This is said to result from the additional litigation costs incurred by both private and public actors in order to enforce subrogation rights. House, *supra* note 53, at 105-06.

⁶⁹ House, *supra* note 53, at 106.

⁷⁰ Hernquist, *supra* note 61, at 177.

⁷¹ Guillermo Gabriel Zorogastua, *Improperly Divorced from Its Roots: The Rationales of the Collateral Source Rule and Their Implications for Medicare and Medicaid Write-Offs*, 55 U. KAN. L. REV. 463, 463 (2007).

⁷² Wershba, *supra* note 55, at 353-54 (noting that these post-verdict hearings likely increase the litigation and administrative costs of actions subject to the collateral source rule).

⁷³ Some states require that an employer repay state unemployment funds if a plaintiff's back pay award is reduced by the amount of unemployment insurance benefits received. See, e.g., TEX. LABOR CODE § 210.001 (2011); MASS. GEN. LAWS ch. 151A, § 69C (2004); CAL. UNEMP. INS. CODE § 1382 (2010). At least one state does not set off unemployment insurance benefits from back pay awards, but allows back pay to be considered employment such that it serves to toll the accrual of unemployment benefits. See TENN. CODE ANN. § 50-7-303(e) (2008).

giving the plaintiff a reduced award.⁷⁴ Thus, in states with subrogation statutes, a prevailing plaintiff who is awarded back pay damages for a period in which that plaintiff also collected unemployment insurance benefits would typically receive as his back pay damages the difference between his total lost wages and the unemployment benefits he had received during the benefit period. The defendant-employer is then required to directly remit to the unemployment insurance fund an amount equal to the overlapping benefits.⁷⁵

The following table illustrates the effect of subrogation, using the same income, back pay damages, and unemployment insurance compensation figures as Table 1:

⁷⁴ See, e.g., COLO. REV. STAT. ANN. § 8-73-110 (2011) (“[A]n individual who has an award for any week and for which week he, at a subsequent date, received a pay award by reason of a decision of the national labor relations board or other source, as a result of the action taken by the National Labor Relations Board or other source, shall immediately repay to the division such amounts as will reimburse the division for all benefit payments made for the period during which he drew benefits and for which the national labor relations board or other source has caused a payment to be made in the form of back pay award to the claimant; and the employer's account charged for such benefits shall be credited accordingly.”); see also ALA. CODE § 25-4-78 (LexisNexis 2011); DEL. CODE ANN. tit. 19, § 3325 (2005); 820 ILL. COMP. STAT. ANN. 405/900 (West 2011); IND. CODE ANN. § 22-4-13-1 (LexisNexis 2011); IOWA CODE ANN. § 96.3 (2011); KAN. STAT. ANN. § 44-719 (2000); KY. REV. STAT. ANN. § 341.415 (West 2011); MD. CODE ANN., LAB. & EMPL. § 8-809 (LexisNexis 2011); MINN. STAT. ANN. § 268.085 (West 2010); MO. REV. STAT. § 288.381 (West 2005); NEV. REV. STAT. ANN. § 612.371 (LexisNexis 2011); 43 PA. STAT. ANN. § 874 (2011); VA. CODE ANN. § 60.2-634 (2010); WASH. REV. CODE ANN. § 50.20.190 (West 2002); WYO. STAT. ANN. § 27-3-306 (2011).

⁷⁵ Though this Note refers to statutes having this effect as “subrogation statutes,” they may share some characteristics with repayment arrangements (agreements by which a victim agrees to pay back an insurer for benefits received when he sues and collects from an injurer). See SHAVELL, *supra* note 54, at 238-39.

In traditional insurance relationships, it may be that a repayment arrangement would be a disincentive to bringing suit, since most or all of the award would necessarily be repaid to the insurer. *Id.* at 239. This does not likely hold true in the case of repayment of unemployment insurance benefits in Title VII firing suits, since those benefits are typically 50% or less of a claimant’s salary.

Table 2 – Subrogation and Back Pay		
	With subrogation statute	Without subrogation statute
Back pay award	\$6,000	\$6,000
Unemployment benefits received	\$2,772	\$2,772
Amount paid by Defendant to Plaintiff	\$3,228	\$6,000
Amount reimbursed to unemployment insurance fund	\$2,772	\$0
Total damages paid by Defendant	\$6,000	\$6,000
Change in unemployment insurance fund balance	\$0	(\$2,272)

As demonstrated in Table 2, subrogation ensures both that the plaintiff receives the “correct” amount of compensation for loss of wages and that the unemployment insurance fund balance remains as if a discriminatory firing had not occurred. Moreover, the employer is still responsible for the full cost of his discriminatory firing. It is also evident that without subrogation, with the plaintiff receiving both unemployment insurance benefits and a back pay award, the assets of the unemployment insurance fund are impaired.⁷⁶

Though subrogation is the most common statutory approach to the problem discussed in this Note, the majority of states do not vest any legal subrogation right in their unemployment insurance funds. In these states, prevailing Title VII plaintiffs receive both a full back pay award and unemployment insurance benefits absent judicial intervention. As described earlier in this Note, many of such states are located in federal circuits that follow the restrictive rule, where district court judges cannot consider unemployment insurance benefits received by plaintiffs and, as a result, the

⁷⁶ Though, of course, this impairment is not necessarily a problem: unemployment insurance funds exist to pay out benefits to the unemployed.

plaintiff receives both the unemployment insurance benefits and the full back pay award.

III. APPROACH BY THE CIRCUIT COURTS OF APPEALS

The Supreme Court has simply never resolved the question presented by this Note. The most frequently cited Supreme Court case in this area is *NLRB v. Gullet Gin*, in which the Court held that the National Labor Relations Board did not abuse its discretion when it refused to reduce a settlement under the National Labor Relations Act by amounts received as unemployment insurance benefits; but even there, the Court made no affirmative holding on the Board's discretion to make such a deduction.⁷⁷ The Court stated only in dicta that unemployment insurance benefits were collateral sources, on the basis that the state, not the employer, made such payments, and because the NLRB had a long-standing practice of refusing to deduct such benefits.⁷⁸

The *Gullet Gin* ruling has failed, however, to elucidate this area of the law. It has been interpreted both as supporting the discretion to deduct unemployment insurance benefits, since it upheld the NLRB's discretionary approach to withholding,⁷⁹ and as prohibiting discretion by its dicta regarding the collateral source rule.⁸⁰ In the absence of a clear directive from the Supreme Court, the federal circuit courts of appeals remain split as to whether or not district courts are prohibited from considering unemployment benefits when calculating back pay awards, or whether the those courts may, in their discretion, reduce back pay awards by unemployment insurance benefits received. This Part examines each approach in turn.

⁷⁷ *NLRB v. Gullet Gin Co.*, 340 U.S. 361, 364 (1951).

⁷⁸ *Id.* at 365-66.

⁷⁹ *See, e.g., EEOC v. Fin. Assur., Inc.*, 624 F. Supp. 686, 694 (W.D. Mo. 1985) (quoting *Gullet Gin*, and stating that, "by analogy, one might well argue that a similar discretion—either to deduct or to refuse to deduct—is vested in the courts in connection with administering Title VII.").

⁸⁰ *See, e.g., Brown v. A.J. Gerrard Manu. Co.*, 715 F.2d 1549, 1550-51 (11th Cir. 1983).

A. THE RESTRICTIVE RULE: UNEMPLOYMENT INSURANCE BENEFITS
CANNOT BE CONSIDERED IN CALCULATING BACK PAY AWARDS

The Third,⁸¹ Fourth,⁸² Sixth,⁸³ Eighth,⁸⁴ Ninth,⁸⁵ and Eleventh⁸⁶ Circuits have held that unemployment insurance benefits are collateral sources that courts cannot consider when calculating a plaintiff's back pay damages. The cases rely on the traditional definition and treatment of collateral sources;⁸⁷ legislative intent;⁸⁸ preference for shifting any double recovery to plaintiffs over defendants;⁸⁹ and furthering the statutory objective to end employment discrimination.⁹⁰

Typical is the approach of the Sixth Circuit Court of Appeals in *Thurman v. Yellow Freight Systems, Inc.* That court strongly opposed allowing the district courts discretion to consider unemployment insurance benefits, holding that that such benefits were plainly a collateral source that were paid to serve a social policy of the state, rather than to discharge an obligation of the employer.⁹¹ The court there also based its holding on its

⁸¹ See *Craig v. Y & Y Snacks, Inc.*, 721 F.2d 77, 82 (3d Cir. 1983).

⁸² See *EEOC v. Ford Motor Co.*, 645 F.2d 183, 195-96 (4th Cir. 1981) (reversed on other grounds).

⁸³ See *Thurman v. Yellow Freight Systems, Inc.*, 90 F.3d 1160, 1170-71 (6th Cir. 1996).

⁸⁴ See *Gaworski v. ITT Commercial Fin. Corp.*, 17 F.3d 1104, 1114 (8th Cir. 1994) (noting that the state had a subrogation statute).

⁸⁵ See *Kauffman v. Sidereal Corp.*, 696 F.2d 343 (9th Cir. 1982).

⁸⁶ See *Brown v. A.J. Gerrard Mfg. Co.*, 715 F.2d 1549 (11th Cir. 1983) (reversing prior 11th Circuit precedent allowing the district courts to make such a deduction on the basis that, at the time *Gullet Gin* was decided, it was the NLRB's practice to always refuse to make such deductions, and as such the refusal to deduct had become "settled back pay law" under the NLRA, which served as the model for the Title VII back pay provision).

⁸⁷ See *McKenna v. City of Phila.*, 636 F. Supp. 2d 446, 457 (2009).

⁸⁸ See *Maxfield v. Sinclair Int'l.*, 766 F.2d 788, 793 (3d Cir. 1985) (noting that Congress included a deduction for interim earnings and amounts reasonably earnable, but failed to provide for other setoffs).

⁸⁹ *Craig v. Y & Y Snacks, Inc.*, 721 F.2d 77, 83 (3d Cir. 1983) ("There is no reason why the [unemployment] benefit should be shifted to the defendant, thereby depriving the plaintiff of the advantage it confers.").

⁹⁰ *Id.* at 84.

⁹¹ See *Thurman v. Yellow Freight Sys., Inc.*, 90 F.3d 1160, 1170-71 (6th Cir. 1996).

belief that two “identically situated claimants” could not be made whole by “radically different backpay awards.”⁹²

In accord with *Thurman* was the Eighth Circuit Court of Appeals in *Gaworski v. ITT Commercial Finance Group*, where the court reversed a district court ruling deducting unemployment benefits from a back pay award.⁹³ The court held that the collateral source rule applied even when the employer contributes to the unemployment insurance fund,⁹⁴ and noted that the deterrence purpose of back pay awards was ill-served by deduction of unemployment insurance benefits because it made discrimination less costly for defendant-employers.⁹⁵

Though it found the question “extremely close and one over which reasonable persons could differ,” the Third Circuit Court of Appeals in *Craig v. Y & Y Snacks, Inc.* reversed a district court’s reduction of a back pay award by unemployment benefits received.⁹⁶ The reversal was, in large part, grounded on the *Craig* court’s finding that unemployment insurance benefits were collateral and intended for the benefit of employees, not employers.⁹⁷ *Craig* went further to declare that deductibility should never be left to the discretion of district court judges, relying on the Supreme Court’s holding that the courts of appeals must apply the back pay provision in a “consistent and principled” manner,⁹⁸ and noting that while back pay might ordinarily be discretionary because it sounds in equity, it has become a presumptive and near-mandatory remedy for prevailing plaintiffs in Title VII suits.⁹⁹ Significantly, however, the court also noted

⁹² *Id.* (quoting *Rasimas v. Mich. Dept. of Mental Health*, 714 F.2d 614 (6th Cir. 1983)). It should be noted that it is not clear that a plaintiff who has received unemployment insurance benefits is identically situated to a plaintiff who has not received said benefits.

⁹³ *See Gaworski*, 17 F.3d at 1114.

⁹⁴ *Id.* at 1112 (quoting *Chi. Great W. Ry. v. Peeler*, 140 F.2d 865, 868 (8th Cir. 1944) (holding that insurance or Workmen’s Compensation Act funds were collateral sources)). The *Gaworski* Court does not indicate whether it believes that the funds are collateral despite employer contribution because: (a) of the incidence of the unemployment insurer tax; (b) direct employee contributions; or (c) regardless of incidence or direct contribution.

⁹⁵ *Id.* at 1113.

⁹⁶ *Craig*, 721 F.2d at 82.

⁹⁷ *See id.* at 82-85.

⁹⁸ *Id.* at 85 (quoting *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 421 (1975)).

⁹⁹ *Id.*

that the plaintiffs in that case were subject to a state subrogation statute,¹⁰⁰ so that their back pay would later be reduced by operation of law.¹⁰¹

B. THE DISCRETIONARY RULE: THE WITHHOLDING OF UNEMPLOYMENT BENEFITS FROM BACK PAY AWARDS IS LEFT TO THE SOUND DISCRETION OF THE DISTRICT COURTS.

The First,¹⁰² Second,¹⁰³ Fifth,¹⁰⁴ Seventh,¹⁰⁵ and Tenth¹⁰⁶ Circuits have held that deduction of unemployment benefits is properly left to the discretion of the district courts. This Part examines both the reasoning for this conclusion and the practices of the district courts in exercising this discretion.

1. Reasoning

Circuits adopting the discretionary approach generally do so with the goal of preventing double recoveries. A robust example is the Second Circuit Court of Appeals in *EEOC v. Enterprise Steamfitters*:

We see no compelling reason for providing the injured party with double recovery for his lost employment; no compelling reason of deterrence or retribution against the responsible party in this case; and we are not in the business of redistributing the wealth beyond the goal of making the victim of discrimination whole.¹⁰⁷

The holding in *Enterprise Steamfitters* was cited favorably and clarified by the Second Circuit in *Dailey v. Société General*. There, the

¹⁰⁰ *Id.* at 83-84.

¹⁰¹ That is, the *Craig* court could not have possibly reduced the plaintiff's back pay award without interfering with the operation of the state unemployment insurance statute or causing a double reduction of back pay.

¹⁰² See *Lussier v. Runyon*, 50 F.3d 1103, 1115 (1st Cir. 1995) (construing Rehabilitation Act of 1973, Pub. L. No. 93-112, 87 Stat. 355 (codified as amended at 29 U.S.C. §§ 701-796i)).

¹⁰³ See *EEOC v. Enter. Ass'n Steamfitters Local No. 638 of U. A.*, 542 F.2d 579, 592 (2d Cir. 1976).

¹⁰⁴ See *Merriweather v. Hercules, Inc.*, 631 F.2d 1161, 1168 (5th Cir. 1980).

¹⁰⁵ See *Bowe v. Colgate-Palmolive Co.*, 416 F.2d 711, 721 (7th Cir. 1969).

¹⁰⁶ See *EEOC v. Sandia Corp.*, 639 F.2d 600, 639 (10th Cir. 1980).

¹⁰⁷ *Enter. Ass'n Steamfitters*, 542 F.2d at 592.

court declined to mandate that unemployment funds be deducted from back pay awards, acknowledging “compelling reasons” for why such benefits should not be deducted,¹⁰⁸ but ultimately left the deduction of said benefits to the “sound discretion” of the district courts instead of following the restrictive rule.¹⁰⁹

Similarly, the Third Circuit Court of Appeals in *Ostapowicz v. Johnson Bronze Co.* upheld a district court decision reducing a plaintiff’s back pay award by the amount of unemployment compensation that either was or reasonably could have been received by the plaintiff, holding that the district court’s deduction represented a “conscientious effort to calculate reasonable and equitable awards under conditions which do not allow for absolute precision.”¹¹⁰

2. Discretion in the District Courts

District courts in circuits following the discretionary rule have focused on several factors to determine whether or not to deduct unemployment insurance benefits from back pay awards. Significantly, some courts have recognized the existence of state subrogation statutes and chosen not to deduct insurance benefits when a plaintiff is subject to subrogation, because the plaintiff’s award will later be reduced by operation of law.¹¹¹ In addition to the effect of subrogation, district courts have refused to reduce a plaintiff’s back pay award where, in the particular circumstances of the case, the plaintiff did not receive a windfall,¹¹² and where the court generally preferred awarding windfalls to plaintiffs rather than defendants.¹¹³

¹⁰⁸ *Dailey v. Societe Generale*, 108 F.3d 451, 460-61 (2d Cir. 1997) (giving, as an example, that where the choice lies with awarding either the plaintiff or the defendant a windfall, the windfall should inure to the plaintiff).

¹⁰⁹ *Id.*

¹¹⁰ *Ostapowicz v. Johnson Bronze Co.*, 541 F.2d 394, 401 (3d Cir. 1976).

¹¹¹ *See, e.g., Cooper v. Asplundh Tree Expert Co.*, 836 F.2d 1544, 1555 (10th Cir. 1988) (holding that an offset for unemployment insurance benefits was “particularly inappropriate . . . because, under Colorado law, an employee who receives a back pay award must repay the Colorado Division of Employment and Training all unemployment benefit payments received for the period covered by the back pay award”).

¹¹² *See Brooks v. Fonda-Fultonville Cent. Sch. Dist.*, 938 F. Supp. 1094, 1110 (N.D.N.Y. 1996).

¹¹³ *See Shannon v. Fireman’s Fund Ins. Co.*, 156 F. Supp. 2d 279 (S.D.N.Y. 2006).

Those district courts that have chosen to reduce back pay awards to account for unemployment insurance benefits have done so for a variety of reasons: because unemployment insurance benefits are not a collateral source;¹¹⁴ because the purpose of the back pay remedy is not to punish employers or to provide windfalls for employees, but rather solely to compensate for a plaintiff's actual economic losses;¹¹⁵ because deduction would have a negligible effect on deterrence;¹¹⁶ and, significantly, because unemployment compensation was not recoverable by the unemployment insurance fund because the jurisdiction lacked a subrogation statute, and "a double recovery was not necessary to make [the] plaintiff whole."¹¹⁷

IV. ANALYSIS

This Part concludes that unemployment insurance benefits should not be treated as traditional collateral sources for the purposes of Title VII back pay awards, both as a matter of law and public policy. Next, this Part reaffirms that the restrictive approach has effected to arbitrarily favor certain plaintiffs over others. Finally, this Part concludes by proposing that all circuits adopt the discretionary approach.

A. UNEMPLOYMENT INSURANCE BENEFITS SHOULD NOT BE TREATED AS SOURCES COLLATERAL TO EMPLOYERS.

Unemployment insurance benefits should not be treated as sources collateral to employers for two reasons. First, insurance coverage paid for by an employer cannot be collateral to that employer, even if employees indirectly pay for a portion of coverage. Second, the policy justifications that underlie the collateral source rule in its traditional applications are inconsistent with the nature of back pay relief under Title VII and in the context of unemployment insurance benefits.

¹¹⁴ *Truskoski v. ESPN, Inc.*, 823 F. Supp. 1007, 1015 (D. Conn. 1993) ("While collateral sources are not offset, unemployment compensation is not from a source independent of the employer . . . [m]aking a person discriminated against whole is not achieved by awarding damages in excess of the actual loss when the excess does not come from a collateral source.").

¹¹⁵ *See Cole v. Uni-Marts, Inc.*, 88 F. Supp. 2d 67, 76 (W.D.N.Y. 2000).

¹¹⁶ *See Wilcox v. Stratton Lumber Co.*, 921 F. Supp. 837, 843 (D. Me. 1996).

¹¹⁷ *Thurber v. Jack Reilly's Inc.*, 521 F. Supp. 238, 243 (D. Mass. 1981) ("[E]quitable considerations militate in favor of a reduction of the gross back pay award here [where] [t]he unemployment compensation paid to the plaintiff is not recoverable from her by the commonwealth.").

1. Unemployment Insurance Benefits Are Not Collateral Sources as Traditionally Defined

Collateral benefits are “compensation . . . from a source independent of the tortfeasor,”¹¹⁸ that is, a collateral source is one “other than the injurer.”¹¹⁹ In the context of Title VII, however, it is the employer who is the injurer. And though unemployment insurance benefits are actually paid out by the government, to the extent that the *source* of compensation is an employer, these payments cannot be considered “collateral.”

True, state governments act as administrators of unemployment insurance funds and collect premiums through taxation. But relying on this aspect of the unemployment insurance relationship to characterize unemployment benefits as collateral privileges form over function. First, the state’s role as intermediary has no bearing on the fact that employers are the sole direct source of funding that provides for the unemployment insurance benefits. Indeed, we would consider ordinary insurance premiums paid by an individual insured to be sourced from that individual, even though the premiums are later intermingled with other insureds’ premiums and invested by an insurer, as the states similarly do with employer unemployment insurance tax proceeds. Second, insurance relationships are nearly always characterized by the presence of an intermediary—typically, insurance contracts create a principal-agent relationship, with the purchaser of insurance acting as principal, appointing the insurer as his agent to take care of insured losses on his behalf.¹²⁰ Through a slight variation of this familiar agent/principal lens, states are the agents designated by federal and state law to represent the employer-principal and to discharge its obligations to the employee-beneficiaries. This intermediary relationship does not alter the fact that unemployment insurance payments made by the agent-state are still attributable to the principal-employer, and represent the discharge of liability.¹²¹

¹¹⁸ BLACK’S, *supra* note 52.

¹¹⁹ SHAVELL, *supra* note 54, at 142-43.

¹²⁰ See TOM BAKER, INSURANCE LAW AND POLICY 5 (2d ed. 2008).

¹²¹ See RESTATEMENT (THIRD) OF AGENCY § 1.01, cmt. g (2006) (“Employee and nonemployee agents who represent their principal in transactions with third parties on the principal’s account and behalf. Employee-agents whose work does not involve transaction interactions with third parties also act ‘on behalf of’ their employer-principal.”). But note, however, that while the employer/state relationship in the unemployment context may lack the required control element. (“Agency is the fiduciary relationship that arises when . . . the agent shall act on

Of course, as discussed earlier in this Note, the incidence of unemployment insurance taxes may play an important role in resolving this dispute. That is, employees may bear some of the cost of unemployment insurance programs to the extent that the cost of the tax is back-shifted through reduced earnings. But although there is no consensus as to the incidence of the tax, the best estimates find that employees are responsible for no more than a portion of the market tax rate, so that ultimately, employers almost certainly pay for most of the unemployment insurance program. Moreover, it is not clear that unemployment insurance benefits are rendered collateral simply because the employer adjusts employee wages to reflect the cost of unemployment insurance. Even if so, the cost remains shared between the employee and employer, with the average employer responsible for the majority of the cost; the benefits can only colorably be considered collateral to the extent that the employee is responsible for paying for the benefits through a reduced wage. Further, the reasoning for excluding bargained-for insurance coverage as a collateral source—that the insured has paid for the covered event through premiums calculated to that plaintiff’s level of risk—does not hold true when only the base-rate, and not the experience rating, is back-shifted to an employee. In that case, there is no guarantee that the employee has made a contribution proportional to the risk and cost of loss.

With the exception of administration by state governments and, as described above, that risk is *not* related to premium, the unemployment insurance system closely tracks traditional forms of insurance in its operation and structure. Unemployment insurance, like traditional insurance, has as one of its primary functions risk-spreading; in this context, it is attempting to ensure that the risks posed by unemployment to both individuals and society as a whole are spread among employers.¹²² And, as in insurance generally, risk-classification is undertaken by unemployment insurers through the experience rating (in traditional insurance, “underwriting”), in order to charge participants for the amount of risk that they bring to the insurance pool. In addition, the structure of unemployment insurance gestures to concerns about moral hazard and

the principal’s behalf and subject to the principal’s control. . . .” RESTATEMENT (THIRD) OF AGENCY § 1.01), in some instances “relationships that are less than fully consensual and, therefore, not common-law agency relations trigger legal consequences equivalent to those of agency.” *Id.* at cmt. d.

¹²² Tax incidence may, as discussed *infra* Part II.B., distort the effect of this intention.

adverse selection,¹²³ both of which are familiar concepts in ordinary insurance. Though participation in the unemployment insurance system is mandatory, and though unemployment insurance may serve broader social goals, neither is unusual in the insurance context. Many ordinary forms of insurance require mandatory participation,¹²⁴ and have at least a partial function of providing social stability and other positive externalities.¹²⁵

Even if one is inclined to the view that unemployment insurance benefits are payments made as a kind of social welfare, funded, like similar programs, through taxation linked to employment, it is not clear that such payments should be immune from deduction from back pay awards. Judge Richard Posner, for example, posits that such benefits, to the extent that they are financed by the government, should be deducted from back pay awards and that the government should have a right to recovery.¹²⁶ The government, in his view, is another victim (in this case, of a discriminatory firing) and should not alone bear the burden of damages.

Thus, unemployment insurance is most closely associated with the employer, perhaps best analogized as a traditional insurance product that is bought and paid for by employers for the benefit of employees. In the case

¹²³ Policing of insured-side moral hazard – the theoretical tendency of insurance to minimize incentives to protect against or minimize the costs or risks of a loss—is reflected in unemployment insurance requirements that those collecting unemployment benefits actively look for jobs and request their insurance benefit anew each week. (Both of these requirements encourage the unemployed to find new employment, and to stop receiving insurance benefits, faster than they might without the requirements.) Insurer-side moral hazard is less of a concern with unemployment insurance than it is with traditional insurance because unemployment insurance is administered by the states, which lack the profit motive driving insurer-side moral hazard.

Adverse selection – the theoretical tendency for high risk insureds to over-consume insurance and for insurers to screen out high-consumption insureds – is controlled by measures that force employers to “purchase” insurance through mandatory taxation and by “enrolling” all involuntarily unemployed persons who apply for benefits and meet unemployment insurance requirements.

¹²⁴ For example, auto insurance is required in almost every state. BAKER, *supra* note 120, at 451.

¹²⁵ Examples of these positive externalities include liability insurance, which ensures that plaintiffs are guaranteed remuneration for covered losses, thus avoiding an insolvent or unwilling defendant from avoiding responsibility for his bad acts, and property insurance, which is universally required by mortgagors in order to secure their collateral, but which also provides the positive externality of neighborhood stability. *See id.* at 8.

¹²⁶ POSNER, *supra* note 57.

of Title VII, however—where the employer is both the source of funds and the defendant—these funds cannot be considered collateral.

2. Public Policy Does Not Support Extending the Collateral Source Rule to Unemployment Insurance Benefits in the Title VII Context

It is not seriously disputed that application of the collateral source rule usually results in a plaintiff receiving a double recovery, which itself often, but not always, results in a windfall for the plaintiff. This Note has explored many of the various policy rationales that have been used to justify both the initial double recovery and the windfall that may result. But in the particular context of Title VII and unemployment insurance benefits, most of these rationales are simply inapposite.

To start, a double recovery cannot be justified as a means of indirectly awarding punitive damages or providing attorney's fees in Title VII suits. Both punitive damages and attorney's fees are explicitly available in the (relatively recently revised) text of Title VII,¹²⁷ a clear expression of Congress's intent as to how and when fees and damages should be awarded.¹²⁸ Indeed, punitive damages are available on only a limited basis under Title VII;¹²⁹ by so limiting their availability, Congress has indicated it almost surely did not intend punitive damages to be awarded to plaintiffs on the sole (and irrelevant to Title VII) basis of the plaintiff previously having collected unemployment insurance benefits. Attorney's fees, on the other hand, are widely available under Title VII, even (after its 1991 revision) for certain plaintiffs who lose their cases,¹³⁰ and thus it is simply not necessary to account for them by "rounding up" a damages award.

¹²⁷ The Civil Rights Act of 1991, Pub. L. No. 102-166, § 102(b), 105 Stat. 1071, 1073 (1991) (current version at 42 U.S.C. 1981a (2006)).

¹²⁸ See ANDREW S. BURROWS, REMEDIES FOR TORTS AND BREACH OF CONTRACT 162-63 (3d ed. 2005) ("[I]f punishment is desired, it is surely better to administer it through punitive damages [rather than the collateral source rule], where the punishment is explicit and where the amount awarded can be fixed in accordance with the extent to which it is felt the defendant deserves punishment.").

¹²⁹ See *infra* Part II.A.

¹³⁰ See Civil Rights Act of 1964, Pub. L. No. 88-352, § 706(g), 78 Stat. 241, 261 (1964) (current version at 42 U.S.C. § 2000e-5 (2006)) (defendant is liable for attorneys fees if the plaintiff can show that a protected characteristic was considered by the defendant in taking an unlawful employment action, even if the defendant can show that the same decision would have been made without consideration of the protected characteristic, thereby avoiding liability).

Ordinary insurers avoid a double-recovery with the near-universal contractual assertion of a subrogation right, which eliminates the risk of a windfall to the insured by virtue of having obtained a collateral benefit. But, as this Note points out, the majority of states do not hold a statutory subrogation right, and thus have no means of collecting benefits for which there exists an overlap. In those states that have asserted a subrogation right, the concern that employers will not bear the full cost of their wrongdoing has been eliminated, along with the need to choose between awarding a windfall to either the employee or employer, because though unemployment benefits are deducted from a plaintiff's back pay award, they are then remitted to the unemployment fund by the employer.¹³¹ Similarly, deterrence and social responsibility for wrongdoing are not weakened as against the employer, since it remains fully responsible for lost wages. Indeed, subrogation has the effect of preserving the strength of the insurance pool, as compared to windfalls, which have the opposite effect.¹³²

To be sure, in those states lacking subrogation statutes, deducting unemployment benefits from an award favors defendants, because it reduces the amount of damages that he will have to pay the plaintiff and, in that sense, lowers the cost of its discriminatory act. But unlike other applications of the collateral source rule, in which the defendant's reduced award is not mitigated by an associated cost, unemployment insurance administration is designed so that each firing has a commensurate effect on the unemployment insurance tax rate. Thus, there remains a disincentive to taking the wrongful action that, to some degree, mitigates forces impairing the deterrence effect of damages and social cost of a firing to an employer.¹³³

¹³¹ See *infra* Part II.D.

¹³² See *infra* Table 2. It is feasible that this approach would also provide a benefit to some employers in the form of reduced premiums by, as discussed earlier in this Note, increasing the overall strength of the unemployment fund.

¹³³ It is very important to note, however, that this may not compensate for the full costs of *discriminatory* firings on society, the elimination of which is a primary interest of Title VII. Since the true social cost of such firings is not easily calculated, it is difficult or impossible to determine precisely whether or not the experience rating can capture these costs.

Another potential concern regarding the discretionary rule, even in a system with a subrogation right, is that it may create an incentive to settle cases for less than the full cost of the discriminatory act (though at least equal to or more than the amount of a back pay award less any unemployment insurance benefits received). That is, the defendant-employer knows that in a system with subrogation

B. UNEMPLOYMENT BENEFITS SHOULD NOT BE UNIFORMLY
DISALLOWED AS DEDUCTIONS FROM BACK PAY AWARDS.

Because of the divergent approaches taken by both the circuit courts of appeals and of the state legislatures, prevailing plaintiffs in Title VII suits are placed on very different footings solely on the basis of residency. District court judges in at least 20 states are *required* to award plaintiffs what this Note has shown should be considered a windfall.¹³⁴ In many other states, on the other hand, plaintiffs are subject to subrogation statutes or to judicial reduction of their back pay award. The basis of this differing treatment, however—essentially, disharmony in the law—has nothing to do with Title VII's aim of making victims of discrimination whole and ending employment discrimination.

Take one example: the Fourth Circuit, whose jurisdiction includes Virginia and Maryland, which have subrogation statutes,¹³⁵ and North Carolina, South Carolina, and West Virginia, which do not, has adopted the restrictive approach. Suppose that an employer located in Virginia discriminatorily fires employees living in West Virginia, Maryland, and Virginia. Further suppose that the firing causes all of the employees identical loss of pay, and that the employees collect identical amounts of unemployment insurance benefits as a result of the firing.¹³⁶ Upon prevailing in a Title VII suit, the plaintiff-employees would each receive the same amount in back pay damages. Following the suit, however, those plaintiffs living in West Virginia are not subject to subrogation and enjoy both unemployment insurance benefits and a back pay award; their identically situated coworkers residing in Virginia and Maryland, on the

it will remain fully responsible for the loss upon a verdict for the plaintiff-employee. The employee, however, knows that it will only receive the amount of back pay minus the unemployment insurance benefits upon prevailing, and resultantly has an incentive to accept a settlement for anything more than that amount.

Of course, a plaintiff who believes that he or she will be entitled to punitive or special compensatory damages may be less swayed to settle, as may a high-income plaintiff, for whom the deduction of back pay benefits represents a smaller proportion of the overall award.

¹³⁴ That is, in states that lack subrogation statutes that are located in circuits that do not allow the district courts the discretion to deduct unemployment insurance benefits from back pay awards. *See infra* Table 2.

¹³⁵ *See* MD. CODE ANN., LAB. & EMPL. § 8-809 (LexisNexis 2008); VA. CODE ANN. § 60.2-634 (2006).

¹³⁶ They would not, in reality, since they all live in different states.

other hand, repay the unemployment fund and receive the excess of back pay over amounts received as unemployment insurance compensation. As demonstrated in Table 1,¹³⁷ this difference is significant: unreduced awards may be more than 50% higher than that of plaintiffs whose award is reduced by a court or through subrogation.

The result is indefensible in the context of the rule of law in general and of Title VII in particular, which mandates principled and uniform application of the back pay provision.¹³⁸ In fact, while many circuit courts of appeals have gestured to the ideal of uniform application of Title VII when refusing to allow the discretionary approach, denial of discretion has, as illustrated above, produced the opposite effect of widening differences between similarly-situated plaintiffs.

C. THE DEDUCTION OF UNEMPLOYMENT INSURANCE BENEFITS SHOULD BE LEFT TO THE DISTRICT COURTS.

As just described, statutory methods of subrogation are incoherent across states, rendering identical back pay awards drastically different depending on a prevailing plaintiff's state of residence. The restrictive rule, uniformly banning consideration of unemployment compensation when determining back pay awards, ignores and enables this incoherence. The federal district courts, however, are well-situated—both as fact finders and as a relatively localized adjudicative body—to ensure uniformity of back pay awards. The circuit courts of appeals can, and should, accomplish uniformity by adopting the discretionary rule, allowing district courts to consider evidence of unemployment benefits when calculating a Title VII plaintiff's back pay damages. District courts, in turn, should leave unchanged awards for those plaintiffs whose unemployment insurance benefits will be subject to subrogation, but should reduce back pay awards when such benefits will result in a windfall, perhaps to the extent that those benefits are attributable solely to an employer and not to employees through back-shifting, when and if such a calculation can be reliably made.

The abandonment of the restrictive rule is necessary because the present two-tiered system of compensation for Title VII plaintiffs is in conflict with the “consistent and principled application of the [Title VII] backpay provision” required by the Supreme Court.¹³⁹ And ensuring that Title VII is consistently applied by adopting the discretionary approach is

¹³⁷ *Infra* Table 1.

¹³⁸ *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 421-22 (1975).

¹³⁹ *Id.*

unlikely to diminish the attainment of the statute's ultimate goal: to reduce employment discrimination. Defendants do not receive a windfall in those states that have subrogation statutes, because the employer remains responsible for the entire amount awarded by the court; in states without subrogation statutes, the reduced awards likely have at most a minor impact on the deterrent effect of the back pay provision due to the increased costs associated with a firing for which unemployment insurance benefits and Title VII back pay damages are claimed.

In fact, reducing back pay awards is plainly consistent with congressional intent to put plaintiffs in the "position where they would have been were it not for the unlawful discrimination."¹⁴⁰ In a Title VII firing suit, the position that the plaintiff would have been in absent discrimination is employed; were the plaintiff employed, she would not have received unemployment compensation. Awarding a plaintiff the total amount of back pay that she would have received if she were not fired, while refusing to reimburse the plaintiff to the extent that she received a benefit as a result of a firing, is consistent with the make-whole nature of back pay relief under Title VII.

It cannot be ignored that employment discrimination suits are relatively low value, and that the private attorney general model is likely weakened by the lower incentives to sue that may result from reduced back pay awards. This incentive structure, however, was strengthened by the addition of attorney's fees and punitive damages under the 1991 Civil Rights Act, both of which (perhaps unlike the windfalls at issue here) furthered Congress's expressed intent that the back pay provision make plaintiffs whole, but go no further.

The discretionary approach is also in accord with the purposes of unemployment insurance. Recall that unemployment benefits are paid both to sustain individuals and their families during periods of temporary unemployment and to stabilize the economy during periods of high unemployment. The first purpose is fulfilled when the plaintiff is able to access his or her insurance benefits during unemployment, and is not nullified when those benefits are later recouped. At that point, the plaintiff has either found employment or no longer qualifies as "temporarily" unemployed; in either case, the insurance benefits have accomplished their income-flow-smoothing function. Nor does reduction of back pay awards impact the economy-wide purposes of unemployment insurance, which, again, is important during the actual period of volatility, but does not have its stabilizing function impaired when later recouped. On the contrary, the

¹⁴⁰ 118 CONG. REC. 7168 (1972).

restrictive rule, by lowering unemployment insurance fund balances as demonstrated in Table 2, may have the effect of reducing compensation, increasing the duration of unemployment and decreasing labor force participation.¹⁴¹

Allowing the federal district courts to consider the amount of unemployment insurance benefits a prevailing plaintiff has received will help to ensure that Title VII back pay awards are truly compensatory. While in some cases a discretionary approach may result in some benefit to defendant-employers, in all cases it will ensure that the back pay provision is consistently applied in accord with congressional and Supreme Court mandates.

V. CONCLUSION

The nature of unemployment insurance and the text of Title VII counsel that the circuit courts of appeals allow the district courts the discretion to consider unemployment insurance benefits when calculating back pay awards for plaintiffs. In states that have subrogation statutes, the district courts should impose no offset, but rather allow reimbursement of the state fund by operation of law. This result ensures plaintiffs receive their entitled make-whole relief, holds defendant-employers liable for the full costs of their discriminatory acts, and benefits the entire pool of insureds by not contributing to the further destabilization of state unemployment insurance funds. Even in states that do not assert subrogation rights, consistent application of Title VII suggests that unemployment insurance benefits should be offset from a prevailing plaintiff's back pay award.

Those states that do not have subrogation statutes should consider their role as a large-scale insurer and act to pass laws that guarantee fair and appropriate benefit payouts while considering the rights of all participating insureds to a stable and fairly administered fund. As 16 states have realized, the best way to accomplish this is, like the majority of conventional private insurers, vesting the unemployment insurance funds with a subrogation right.

¹⁴¹ See Anderson, *supra* note 46. As discussed earlier in this Note, market rates for the unemployment tax are calculated by reference to the balance, income, and expenses of the unemployment insurance fund.

LESSONS FROM THE PRICE-ANDERSON NUCLEAR INDUSTRY INDEMNITY ACT FOR FUTURE CLEAN ENERGY COMPENSATORY MODELS

TAYLOR MEEHAN*

The following note discusses the Price-Anderson Nuclear Industry Indemnity Act as a model liability insurance system for future clean energy technologies and sources such as carbon sequestration and geothermal energy. The Price-Anderson Act implements a tiered insurance system that requires individual commercial nuclear power plants to secure private insurance policies for site-specific incidents up to a certain threshold. This first layer of indemnity is then supplemented by an industry-wide pooling system that provides indemnification in the event of an incident that accrues greater financial losses than the initial, primary insurance layer obtained by the responsible nuclear plant. In the event the industry-wide insurance pool funds are exhausted, the federal government is the final indemnifier, providing additional compensation to affected individuals when deemed appropriate. This note considers the history of the Price-Anderson Act, its development and subsequent amendments since its enactment in 1957, and highlights the specific aspects of the system that should be adopted in the future. In particular, the note argues that carbon sequestration technology and geothermal energy are presently situated in a similar situation as the nuclear industry was in the early 1950s. The parallels between the industries – most notably the low risk of an industrial accident, yet extensive consequences in the event of an incident – invite comparison and analysis into whether the nuclear industry indemnity system is a transferable model to future clean energy technologies. Ultimately, the note argues that a number of the key components of the Price-Anderson Act – particularly its liability cap, federal involvement, no-fault liability, federal jurisdiction, and continually written policies – not only are suitable for future systems, but in fact should be implemented by the insurance industry when underwriting the carbon sequestration and geothermal energy insurance system. The note concludes that the United States is in dire need of restructuring its national energy policy and an essential aspect to this national policy is creating an underlying system of liability that can be applicable, with specific adaptations in lieu of inherent

* University of Connecticut School of Law, Juris Doctor Candidate for the Class of 2012.

differences in these technologies, to new clean energy sources. The current system of the Price-Anderson Nuclear Industry Indemnity Act is America's best solution.

I. INTRODUCTION

With the increasing energy demand in the United States and the diminishing supply of traditional domestic fossil fuels,¹ the nation is confronted with serious energy concerns that necessitate a review of our national energy policy. Not only is the United States the largest energy producer, consumer and net importer per capita in the world,² but the

¹ Energy production from fossil fuels (e.g., coal, oil and natural gas) is expected to continue to dominate U.S. energy production for years to come. Naturally, the continual production of fossil fuel reserves and the increased energy demand within the United States has led to significant concerns of a diminishing domestic supply of such resources. While the energy industry is alarmed at the potential diminution in domestic fossil fuel resources, the United States has witnessed a boom in the exploration, development and early production of natural gas reserves located in deep, shale rock formations around the country. These shale formations contain natural gas reserves that were previously considered inaccessible and uneconomical for energy production. However, with the increasing development and use of a technological drilling process called hydraulic fracturing (fracking) these shale fields are opening up a vast amount of potential for natural gas production within the United States. The practice, however, is highly controversial. The energy industry currently heralds fracking as the answer to U.S. energy needs while environmental groups and legislators are concerned with potential groundwater contamination and increased seismic activity within surrounding drilling areas. The national debate on the practice is presently unfolding. See J. DANIEL ARTHUR, P.E., BRIAN BOHM, P.G. & MARK LAYNE, PH.D., P.E., HYDRAULIC FRACTURING CONSIDERATIONS FOR NATURAL GAS WELLS OF THE MARECELLUS SHALE, Ground Water Protection Council 2008 Annual Forum 7-9 (Sept. 21-24, 2008).

² See *International Energy Statistics*, UNITED STATES ENERGY INFORMATION ADMINISTRATION, INDEPENDENT STATISTICS AND ANALYSIS, <http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=44&pid=44&aid=2> (last visited Aug. 21, 2011). But see *China Overtakes the United States to Become World's Largest Energy Consumer*, INTERNATIONAL ENERGY AGENCY (July 20, 2010), http://www.iea.org/index_info.asp?id=1479 (finding China has recently surpassed the United States in total energy consumption, however the United States remains the largest energy consumer per capita).

country also boasts the world's largest coal reserves – making the United States extremely dependent on fossil fuel energy for short-term and long-term economic growth.³ Concerns over energy security and foreign dependence are exacerbated by scientific and social apprehension surrounding the leakage of greenhouse gas from these fossil fuel energy sources into the environment.⁴

Accordingly, the United States must continue to focus on its development of clean energy sources that can help mitigate many of the risks and problems associated with fossil fuel energy. Headlining these developments are clean energy sources and technologies such as nuclear energy, carbon sequestration and geothermal energy. In order to properly support the clean energy movement, the United States must attract significant financial investment from the private sector as well as provide a system of adequate insurance coverage in order to mitigate any associated risks.

The following commentary examines the benefits and deficiencies of the current financial protection program of the nuclear energy industry, which was established by the Price-Anderson Nuclear Industry Indemnity Act. It continues, arguing in support of the development of a similar, yet varied version of the underlying nuclear industry indemnity system to insure future clean energy sources and technologies.

Analysis of the Price-Anderson Act reveals a variety of important issues and concerns for insurers attempting to provide coverage for clean energy technologies and sources such as carbon sequestration and geothermal energy. While these technologies are each distinct and consist of specific, technical issues that are unique to their own field, the Price-Anderson Act offers a general model of a public-private partnership that has successfully insured the nuclear industry for over fifty years.

The following commentary will begin by discussing the history and development of the Price-Anderson system, as it has been amended and renewed four times since its original enactment. The note will then provide analysis of the long-term liability issues associated with the nuclear

³ See Guri Bang, *Energy Security and Climate Change Concerns: Triggers for Energy Policy Change in the United States?*, 38 ENERGY POL'Y 1645, 1645 (2010).

⁴ KEVIN A. BAUMERT, TIMOTHY HERZOG & JONATHAN PERSHING, NAVIGATING THE NUMBERS: GREENHOUSE GAS DATA AND INTERNATIONAL CLIMATE POLICY 12 (World Resources Institute 2005) (The United States is among the leading emitters of greenhouse gases in the world, most notably carbon dioxide. In 2000, the United States amounted for 20.6% of the world's greenhouse gas emissions.).

industry, carbon sequestration and geothermal energy. In particular, the note will argue in support of the establishment of a similar, private-public tiered insurance pool system for future clean energy industries.

II. BACKGROUND TO THE PRICE-ANDERSON ACT

The United States is in a similar position today as it was in the 1950s. In the 1950s, the nation was confronted with the harsh realities of the aftermath of World War II and the increasing industrial growth of the nation. The need for a rise and diversification in its energy production to meet the demand was essential. Therefore, the federal government encouraged energy diversification, invested in research and development of alternative energy sources, and provided regulatory incentives to advance oil, coal and nuclear development within the private sector.⁵

In 1954, Congress passed the Atomic Energy Act, which provided for the development and regulation of civilian and military uses of nuclear materials in the United States.⁶ The Act marked the first time the private sector was encouraged to become a player in the development of commercial nuclear power plants. The initial version of the statute, however, did not establish a system of indemnification for, or limits on, private licensee liability in case of offsite injury to individuals or damage to property.⁷

Thus, the private sector approached the invitation with both caution and uncertainty. The private sector was concerned with the lack of nuclear experience – not only from a technological standpoint but also from an insurance perspective. The lack of certainty prompted a resistance from insurance companies to provide commercial liability coverage for private

⁵ See Roger H. Bezdek & Robert M. Wendling, *A Half Century of US Federal Government Energy Incentives: Value, Distribution, and Policy Implications*, 27 INT'L J. GLOBAL ENERGY ISSUES 42, 43 (2007).

⁶ Atomic Energy Act of 1954, 42 U.S.C. § 2011 (2006) (The Act is the fundamental U.S. law on both the civilian and the military uses of nuclear materials. It provides for the development and regulation of nuclear materials and facilities in the United States. The Act declares that "the development, use, and control of atomic energy shall be directed so as to promote world peace, improve the general welfare, increase the standard of living, and strengthen free competition in private enterprise.").

⁷ Atomic Energy Act of 1954, Pub. L. No. 83-703, 68 Stat. 919 (1954), *amended by* the Price-Anderson Act, Pub. L. No. 85-256, 71 Stat. 576 (1957) (current version at 42 U.S.C. § 2012(i) (2006)).

sector nuclear development.⁸ Accordingly, representatives from the private sector stressed to Congress that they would be forced to withdraw from the field if their liability was not limited by legislation.⁹

III. THE PRICE-ANDERSON NUCLEAR INDUSTRY INDEMNITY ACT

In response to such concerns, Congress passed the Price-Anderson Act in 1957 as an amendment to the Atomic Energy Act. The Price-Anderson Act established a nuclear liability indemnity system and encouraged further development of the nuclear industry within America.¹⁰ This system included a liability cap in the event of a nuclear incident – a provision that was necessary for initiating private investment and development of nuclear energy within the United States.¹¹

The nuclear industry is an area in which large amounts of energy production is accompanied with low, yet devastating, potentials of risk, especially during early developments. Thus, the Price-Anderson Act sought

⁸ See Barry Brownstein, *The Price-Anderson Act: Is It Consistent with a Sound Energy Policy?*, CATO INSTITUTE (Apr. 17, 1984), http://www.cato.org/pub_display.php?pub_id=902 (“Consider the following statements from the 1956 and 1957 hearings on the then-proposed Price-Anderson amendment. A vice president of Westinghouse, Charles Weaver, stated: ‘Obviously we cannot risk the financial stability of our company for a relatively small project no matter how important it is to the country’s reactor development effort, if it could result in a major liability in relation to our assets.’”) (quoting Joint Committee on Atomic Energy, Governmental Indemnity for Private Licensees and AEC Contractors Against Reactor Hazards-Hearings Before the Joint Committee on Atomic Energy, 84th Cong., 2d sess., 1956, p. 110); see also *id.* (“General Electric also indicated during the hearings that it was prepared to halt its work in the nuclear industry should a limitation on liability not be passed.”) (citing Joint Committee on Atomic Energy, Hearings Before the Joint Committee on Atomic Energy on Governmental Indemnity and Reactor Safety, 85th Cong., 1st sess., 1957, p. 148); and *id.* (“Suppliers of reactor shields also indicated their unwillingness ‘to undertake contracts in this field without being relieved of uninsurable liability in some way.’”) (quoting Joint Committee on Atomic Energy, Hearings Before the Joint Committee on Atomic Energy on Governmental Indemnity and Reactor Safety, 85th Cong., 1st sess., 1957, p. 148).

⁹ See *Duke Power Co. v. Carolina Environmental Study Group, Inc.*, 438 U.S. 59, 64 (1978).

¹⁰ Act of September 2, 1957, Pub. L. No. 85-256, 71 Stat. 576 (codified as amended in scattered sections of 42 U.S.C.).

¹¹ 42 U.S.C. §§ 2210(b)(4)(A)(i)-(ii) (2006).

to implement a sufficient liability and compensation framework to both protect the American public in the event of a nuclear incident as well as advance financial investment and development of the industry.¹² The Act is essentially an insurance program that encourages private development of nuclear power, establishes a legal framework for handling potential liability claims, and provides a ready source of funds to compensate injured victims of nuclear accidents.¹³

In drafting the indemnity plan, Congress initially established a two-tiered insurance system. The primary layer of the system required each commercial nuclear power plant to secure its own insurance coverage up to a certain threshold. In the event that the primary layer was exhausted, the federal government would provide an additional layer of financial protection.¹⁴

The initial two-tiered model has since been bolstered to include an additional industry-wide pool that requires nuclear reactors to collectively contribute to a separate insurance pool.¹⁵ Accordingly, the current system consists of a three-tiered system. To date, the primary layer requires each nuclear plant to secure \$375 million in financial protection.¹⁶ In the event of an incident exceeding the primary layer's coverage, the industry-wide pool kicks in and each reactor is assessed a prorated share of the excess up to \$111.9 million.¹⁷ The \$111.9 million is adjusted every five years for

¹² The original act implemented a system that would last for ten years. This was an attempt by legislators and nuclear industry actors to readdress the Amendment once significant development within the nuclear field and commercial liability industry could occur. 42 U.S.C. § 2210 (2006).

¹³ *National Energy Issues: Hearings Before the S. Comm. on Energy and Nat. Resources*, 107th Cong. 53, 54 (2001) (statement of John L. Quattrocchi, Senior Vice President, Underwriting, American Nuclear Insurers, West Hartford, CT) [hereinafter Quattrocchi].

¹⁴ *The Price-Anderson Act: Background Information*, AM. NUCLEAR SOC'Y (Nov. 2005), <http://www.new.ans.org/pi/ps/docs/ps54-bi.pdf>.

¹⁵ The Insurance Institute defines an insurance pool as "a group of insurance companies that pool assets, enabling them to provide an amount of insurance substantially more than can be provided by individual companies to insure large risks such as nuclear power stations." See *Insurance Pools Definition*, INSURANCE INFORMATION INSTITUTE, <http://www2.iii.org/glossary/i/> (last visited Aug. 22, 2011).

¹⁶ *Fact Sheet on Nuclear Insurance and Disaster Relief Funds—Nuclear Insurance: Price-Anderson Act*, UNITED STATES NUCLEAR REGULATORY COMMISSION (Jun. 2011), <http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/funds-fs.html>.

¹⁷ *Id.*

inflation and represents the maximum retrospective assessment that each insured licensee can be assigned per incident.¹⁸

The additional pool, also known as the Secondary Financial Protection program, is currently comprised of 104 power reactors and amounts to nearly \$12.6 billion dollars.¹⁹ This industry-wide retrospective rating program will be used in the event that a loss exceeds the primary insurance limit.²⁰ In turn, if the second tier is fully exhausted, Congress is committed to determine whether additional relief is needed.²¹ If Congress determines additional relief is necessary, the federal government is the final indemnifier.²²

Since its enactment, the Price-Anderson Act has been amended in 1966, 1975, and 1988. Recently, the Act was renewed with the passage of the Energy Policy Act of 2005, which extended the program until December 31, 2025.²³ The following sections discuss the development and amendments to the Price-Anderson Act and their significance to the legal and insurance framework of the nuclear industry.

A. 1957 PRICE-ANDERSON ACT

In 1957, Dwight Eisenhower signed into law the Price-Anderson Act, establishing the first nuclear indemnification plan for commercial nuclear power plants within the United States.²⁴ The Act initially required a commercial nuclear power plant licensee with energy capacities of 100,000 electrical kilowatts or more to obtain \$60 million of financial protection –

¹⁸ See Quattrocchi, *supra* note 13, at 56.

¹⁹ *Need for Nuclear Liability Insurance*, AMERICAN NUCLEAR INSURERS 2 (Jul. 2011), <http://www.amnucins.com/library/Nuclear%20Liability%20in%20the%20US.pdf> [hereinafter ANI Liability Insurance].

²⁰ See Quattrocchi, *supra* note 13, at 59; see also 42 U.S.C. § 2014(k) (2006).

²¹ 42 U.S.C. § 2210(e)(2) (2006).

²² 42 U.S.C. § 2210(c) (2006). (The federal indemnity agreement covers liability for any bodily injury, sickness, disease or death, or loss of or damage to property by a nuclear incident occurring within the United States. See 42 U.S.C. § 2214(q) (2006)).

²³ Cole Mahone Adams, *Damages and Injury: Smith v. Carbide and Chemicals Corporation and the Application of Kentucky Law under the Price-Anderson Act*, 22 J. NAT. RESOURCES & ENVTL. L. 175, 177 (2008-2009).

²⁴ See S. REP. NO. 85-296, at 8 (1957) (The Act provides the United States with “a practical approach to the necessity of providing adequate protection against liability arising from atomic hazards, as well as a sound basis for compensating the public for any possible injury or damage arising from such hazards.”).

the maximum amount of private insurance potentially available at the time – in order to remain in operation.²⁵ In the event of a nuclear incident, the Atomic Energy Commission agreed to indemnify the nuclear operators or manufacturers for all liability up to, but not in excess of, \$500 million.²⁶

The initial act established the precedent for a liability cap for the federal government and also included “omnibus coverage,” which extended coverage not only to a person with whom an agreement of indemnification was executed but also to any person or persons deemed liable under state tort law.²⁷ While the original \$500 million proved to be a rough estimation of liability, each successive amendment Congress has addressed, and raised, the liability cap to reflect an appropriate balance between industry capacity and potential harm.²⁸

B. PRICE-ANDERSON ACT: 1966 AND 1975 AMENDMENTS

In accordance to the 1957 version of the Price-Anderson Act, the statute was to expire following a ten-year trial period.²⁹ Congress, however, extended the bill in 1966³⁰ and again in 1975.³¹ The 1966 amendment addressed three major concerns of legal impediments claimants faced when seeking relief under the Act – proving legal causation, state statutes of limitations and jurisdictional variances.

The Joint Committee tasked to remedy the deficiencies of the Act was concerned that the burden of establishing causation was too stringent, as many state tort laws required findings of fault or negligence.³² The argument followed that proving the fault or negligence standard was too difficult of a burden on the individual victim. Thus, in order to address the uncertainty in state tort law regarding the applicability of causation, the

²⁵ Price-Anderson Act, Pub. L. No. 85-256, 71 Stat. 577 (1957) (codified as 42 U.S.C. § 2210(b) (2006)).

²⁶ Dan M. Berkovitz., *Price-Anderson Act: Model Compensation Legislation?—The Sixty-Three Million Dollar Question*, 13 HARV. ENVTL. L. REV. 1, 7 (1989).

²⁷ *Id.* at 8.

²⁸ See Quattrocchi, *supra* note 13, at 58.

²⁹ See Price-Anderson Act, Pub. L. No. 85-256, 71 Stat. 576 (1957) (current version at 42 U.S.C. § 2011 (2006)).

³⁰ See Act of September 29, 1965, Pub. L. No. 89-210, 79 Stat. 855.

³¹ See Act of December 31, 1975, Pub. L. No. 94-197, 89 Stat. 1111.

³² David M. Rocchio, *The Price-Anderson Act: Allocation of the Extraordinary Risk of Nuclear Generated Electricity: A Model Punitive Damage Provision*, 14 B.C. ENVTL. AFF. L. REV. 521, 538 (1987).

1966 amendments included a provision for the waiver of various defenses under state tort law in the event of a major accident termed an “extraordinary nuclear occurrence.”³³ This provision was enacted in order to assure that the victim’s entitlement to compensation would be determined under a strict liability standard, instead of the negligence standard that most state courts require.³⁴

In addition, the Committee addressed the fact that due to the latent nature of injury, harm and damage caused by exposure to radioactive material, state statutes of limitation would most likely invalidate any claims as untimely.³⁵ As a result, the 1966 Amendment provided a provision that waived the application of state statutes of limitations that were more restrictive than the three-year limit specified by the Act.³⁶ Finally, the 1966 amendment invoked a removal provision, which brought claims arising out of an extraordinary nuclear occurrence within the jurisdiction of federal district courts.³⁷ All claims resulting from the same “extraordinary nuclear occurrence” were to be consolidated into one federal court. The court would then be responsible for adjudicating all claims, distributing any compensatory damages if necessary and prioritizing any payouts in the event of fiduciary exhaustion.³⁸

In 1975, Congress reauthorized the Act through 1987. The 1975 amendments drastically changed the system by beginning to phase out the \$500 million layer of federal indemnity. The amendment shifted the secondary layer of protection instead to the nuclear industry and private

³³ 42 U.S.C. § 2014(j) (1982) (An extraordinary nuclear occurrence is defined “any event causing a discharge or dispersal of source, special nuclear, or byproduct material from its intended place of confinement in amounts offsite, or causing radiation levels offsite, which the Nuclear Regulatory Commission or the Secretary of Energy, as appropriate, determines to be substantial, and which the Nuclear Regulatory Commission or the Secretary of Energy, as appropriate, determines has resulted or will probably result in substantial damages to persons offsite or property offsite.” When determining whether an incident is to be considered an extraordinary nuclear occurrence, the Nuclear Regulatory Commission established a set of criteria that can be found in 10 C.F.R. §§ 140.81-140.85 (1988)).

³⁴ See S. REP. NO. 89-1605, at 3-4 (1966).

³⁵ See Rocchio, *supra* note 32, at 525.

³⁶ See 42 U.S.C.A. § 2210(n)(1)(F)(iii) (2006) (The Act allows “any issue or defense based on any statute of limitations if suit is instituted within three years from the date on which the claimant first knew, or reasonably could have known, of his injury or damage and the cause thereof.”).

³⁷ 42 U.S.C. § 2210(n)(2) (2006).

³⁸ 42 U.S.C. § 2210(n)(3)(A)-(C) (2006).

insurance companies.³⁹ The Act required each nuclear plant to contribute up to \$5 million of retrospective premiums in the event of a nuclear accident at any commercial nuclear plant within the United States for which damages exceeded the required \$60 million amount of private insurance for each site.⁴⁰ The total amount of financial protection in this secondary layer depended on the number of operating power plants, however the government retained the assurance that it would provide compensation in the event that the total protection was less than the previous amount of \$560 million.⁴¹

C. 1988 AMENDMENTS

In the aftermath of the 1979 accident at the Three Mile Island nuclear power plant,⁴² Congress opted to further increase the liability cap and financial protections of the Act.⁴³ In the 1988 amendment, Congress increased the liability of the nuclear industry to \$9.87 billion dollars, nearly ten times greater than the original liability cap.⁴⁴

Following the Three Mile Island incident, lawsuits were filed in state and federal courts due to the language of the Act – that is, only “extraordinary nuclear occurrences” could be consolidated in federal court.⁴⁵ Thus, Congress amended the Act in 1988 by granting United States district courts with original removal jurisdiction over all “public liability

³⁹ Berkovitz, *supra* note 26, at 14.

⁴⁰ *Id.* at 14-15.

⁴¹ *Id.* at 15.

⁴² See Nuclear Regulatory Commission. *Three Mile Island Accident: Backgrounder*, NUCLEAR REGULATORY COMMISSION, available at <http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html> (last visited Aug. 22, 2011) (On March 28, 1979, the Three Mile Island nuclear power plant near Middletown, Pa., suffered a severe core meltdown, leading to the most serious nuclear incident in U.S. commercial nuclear power plant operating history. No deaths or injuries to plant workers or members of the nearby community occurred, but it brought widespread change to the security, operation, emergency response and regulations of the nuclear industry).

⁴³ *Id.*

⁴⁴ Berkovitz, *supra* note 26, at 41.

⁴⁵ See *El Paso Natural Gas Co. v. Neztosie*, 526 U.S. 473, 486 (1999) (The Supreme Court held, among other issues, that the Price-Anderson Act’s terms “are underscored by its legislative history, which expressly refers to the multitude of separate cases brought ‘in various state and Federal courts’ in the aftermath of the Three Mile Island accident.”).

actions” arising under the Price-Anderson Act.⁴⁶ This amendment, combined with the waiver of defense provisions, the omnibus coverage and the predetermined sources of funding, provided individuals seeking legal recourse significant advantages in federal court that might not otherwise be offered under state tort law.⁴⁷ The substantive rules for decision, however, remain derived from state law in which the nuclear incident occurs, unless such law is inconsistent with the provisions of the Price-Anderson Act.⁴⁸

D. PRICE-ANDERSON ACT SINCE 2005

With strong bipartisan support, Congress passed the Energy Policy Act (EPA) of 2005, which, among other provisions, provided for the extension of the Price-Anderson Act from 2005 until December 31, 2025.⁴⁹ This is the longest extension of the program since its enactment. The most significant amendment from EPA is the increase in the amount of annual financial contributions from commercial reactors.⁵⁰ The Act now requires individual site operators to provide \$375 million of primary financial protection and to contribute \$111.9 million to the Secondary Financial Protection Program, plus 5% for legal costs per reactor.⁵¹

IV. THE REGULATING AGENCY: THE NUCLEAR REGULATORY COMMITTEE

Congress established the Nuclear Regulatory Commission (NRC) as an independent regulating agency designed to license and regulate the nation’s civilian use of nuclear materials to ensure adequate protection of public health and safety, to promote the common defense and security, and to protect the environment.⁵² An essential component of the NRC’s role as

⁴⁶ 42 U. S. C. § 2210(n)(2) (2006).

⁴⁷ *O’Conner v. Commonwealth Edison Co.*, 13 F.3d 1090, 1100 (7th Cir. 1994) (citing S. REP. NO. 100-218, at 4 (1987)).

⁴⁸ 42 U.S.C. § 2014(hh) (2006).

⁴⁹ Price-Anderson Amendments Act of 2005, Pub. L. No. 109-58, §§ 602-608, 119 Stat. 779, 779-781 (2005).

⁵⁰ 70 FR 61885-01, Rules and Regulations, Nuclear Regulatory Commission (2005) (codified in scattered sections of 10 C.F.R.).

⁵¹ Nuclear Regulatory Commission, *supra* note 16, at 1.

⁵² William C. Ostendorff, Commissioner, Nuclear Regulatory Comm’n, Keynote Address at the Emerging Issues Policy Forum, Powering the Future 2010: Nuclear Regulation and the Nuclear Renaissance 1-2 (Oct. 4, 2010), *available at* <http://pbadupws.nrc.gov/docs/ML1027/ML102790151.pdf>.

a federal agency is that it is not an entity that promotes the use of nuclear and radiation technologies,⁵³ but instead is one that regulates the use of such technologies to ensure the safety and security of the nuclear industry.⁵⁴

As stated by Commissioner William C. Ostendorff during a keynote address on the nuclear renaissance in Amelia, FL,⁵⁵ the agency strives to adhere to its principles of good regulation through independence, transparency, efficiency, clarity, and reliability.⁵⁶ If the nation is to proceed with clean, alternative energies as an integral part of our national energy policy, then the clean energy technologies will also need an independent regulatory agency similar to the NRC. Such an agency would ensure best practice techniques, regulation of licenses and operations, and uniformity across the industry as well as provide direction as an oversight committee in promoting safety and the public interest.

V. THE INSURERS: THE AMERICAN NUCLEAR INSURERS

Since the establishment of the Price-Anderson system, a group of member insurance companies – the American Nuclear Insurers (ANI) – has been responsible for all of the nuclear liability policies.⁵⁷ American Nuclear Insurers is an unincorporated voluntary joint underwriting association that directly writes nuclear liability insurance for nuclear facilities.⁵⁸ In order to be a member company of ANI, insurers are required to contain an A.M.

⁵³ *Id.* at 2. The role of promoting nuclear technologies was assigned to the Department of Energy and its predecessor, the Energy Research and Development Administration (ERDA), in the 1970s.

⁵⁴ *Id.*

⁵⁵ Commissioner Ostendorff is one of the five head members (commissioners) of the Nuclear Regulatory Commission. NUCLEAR REGULATORY COMMISSION, The Commission, <http://www.nrc.gov/about-nrc/organization/commfuncdesc.html> (last visited Sept. 5, 2011).

⁵⁶ See Ostendorff, *supra* note 52, at 2.

⁵⁷ As of January 1, 1998, the insurance pools had underwritten the following policies: Operating power reactors: 69 sites; Non-power reactors: 27; Fuel fabrication facilities: 6; Waste disposal and storage facilities: 12; Miscellaneous facilities including nuclear laundries and research laboratories: 55; Discontinued nuclear facilities: 20; Suppliers and transporters: 225. See Paul Bailey, THE PRICE-ANDERSON ACT - CROSSING THE BRIDGE TO THE NEXT CENTURY: A REPORT TO CONGRESS, ICF INCORPORATED FOR THE NRC 75 (Oct. 1998).

⁵⁸ American Nuclear Insurers, *Overview*, <http://www.amnucins.com/AboutAni.html> (last visited Aug. 21, 2011) [hereinafter ANI].

Best Rating of “A-” or better, possess a policyholder surplus (PHS) of at least \$100 million, and release unqualified, audited financial statements for the latest financial reporting period.⁵⁹ Currently, there are 21 insurance companies that are member companies of ANI.⁶⁰

American Nuclear Insurers manages both domestic and foreign underwriting syndicates.⁶¹ The domestic syndicate provides third party nuclear liability insurance to every commercial nuclear power plant in the United States as well as other entities that support the operation of power plants such as fabricators of nuclear fuel, nuclear research facilities, waste management and disposal facilities, and companies that supply any goods and services to the nuclear industry.⁶²

Under the foreign underwriting syndicate, ANI participates in reinsurance programs in 18 foreign countries.⁶³ ANI retains around a third of the liability exposure under each policy while ceding the remaining amount to insurers around the world.⁶⁴ This approach allows ANI to organize the resources of the worldwide insurance community and spread the uncertainties of the risk over a large financial base.

American Nuclear Insurers provide four specific liability policies in order to satisfy nuclear plant’s requirements under the Price-Anderson Act. These policies include a Facility Form Policy, Secondary Financial Protection Program, Facility Worker Form Policy, and Supplier’s and Transporter’s Policy.⁶⁵ The Facility Form Policy is the site-specific insurance coverage that owners or operators of commercial nuclear power plant are required to have under the Price-Anderson Act.⁶⁶ This coverage is strictly limited to liability for bodily injury or offsite property damage

⁵⁹ Richard Jones, *Nuclear Insurance: Where Does it Fit in the Green Generation?*, 16 J. REINSURANCE 71, 74 (Spring 2009).

⁶⁰ *Id.* at 75. These insurance companies include Ace American Insurance Company, Employers Mutual Casualty Company, AXIS Reinsurance Company, Federal Insurance Company (Chubb), Swiss Re America, and State Farm Mutual Auto Insurance Co., among others.

⁶¹ *Id.* at 76.

⁶² *Id.*

⁶³ *Id.*

⁶⁴ Ian Hore-Lacey, World Nuclear Association, *The Encyclopedia of Earth, Price-Anderson Act of 1957, United States* (Dec. 7, 2009, 10:35 PM), http://www.eoearth.org/article/Price-Anderson_Act_of_1957_United_States (last accessed Aug. 22, 2011).

⁶⁵ See ANI Liability Insurance, *supra* note 19, at 1-4.

⁶⁶ *Id.* at 1-2.

caused by nuclear material.⁶⁷

Underwriters for the Facility Form Policies included two distinct provisions in order to tailor it to the nuclear industry. First, the policies are written on a continuous basis with no explicit end date. The insurance coverage ends only when the owner or operator of the commercial nuclear power plant or ANI cancels the terms.⁶⁸ Any claims resulting from the policy term remains under the coverage. This takes into account the latent nature of any damage or harm stemming from nuclear exposure.⁶⁹ Second, the policies cover not only the owner or operator of the plant but also any entity connected with the nuclear plant, thus assuring all third party nuclear liability claims will be covered while also preventing potential stacking of limits.⁷⁰

The Secondary Financial Protection Program, discussed previously, provides for the industry-wide indemnification in the event of a nuclear incident that exceeds the site-specific insurance policy. The structure of the insurance coverage under the Price-Anderson Act has enabled insurers to provide stable, high quality coverage for nuclear risks.

VI. PRICE-ANDERSON ACT AS MODEL COMPENSATION

The Price-Anderson Act represents the balancing of the interests and needs of the public not only as private citizens but also as consumers in and beneficiaries of the private business enterprise of nuclear energy.⁷¹ The following sections highlight the components of the Act that make it a model compensation system for clean energy technologies that have not yet enjoyed the long history and maturity of the nuclear industry.

A. NECESSITY OF A LIABILITY CAP

In order to encourage development of the nuclear industry as well as provide adequate protection for the American public, Congress implemented the liability cap for the nuclear industry in order to strike the appropriate balance of accountability and development. This limitation, however, does not directly limit the ability of individual claimants affected by any nuclear incident from recovering. As discussed in *Duke Power v.*

⁶⁷ See Jones, *supra* note 59, at 77.

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ *Id.* at 78.

⁷¹ See Quattrocchi, *supra* note 13, at 55.

Carolina Environmental Study Group, Inc., the legislative history of the Act clearly indicates that the primary – and secondary insurance pools – are not figures that were arrived at on the supposition that it alone would be sufficient to guarantee full compensation in the event of a nuclear accident.⁷² The initial primary insurance was conceived of as a “starting point” or “a working hypothesis” derived from expert appraisals of the exceedingly small risk of a nuclear incident involving claims in excess of that figure.⁷³ This figure has risen from \$560 million to \$12.6 billion over the past four decades in order to ensure public protection. In addition, legislative history indicates that Congress would likely enact extraordinary relief provisions in order to provide for additional relief.

[T]his limitation does not, as a practical matter, detract from the public protection afforded by this legislation. In the first place, the likelihood of an accident occurring, which would result in claims exceeding the sum of the financial protection, required and the governmental indemnity is exceedingly remote, albeit theoretically possible. Perhaps more important, in the event of a national disaster of this magnitude, it is obvious that Congress would have to review the problem and take appropriate action. The history of other natural or man-made disasters, such as the Texas City incident,^[74] bear this out. The limitation of liability serves primarily as a device for facilitating further congressional review of such a situation, rather than as an ultimate bar to further relief of

⁷² *Duke Power Co. v. Carolina Env'tl. Study Grp., Inc.*, 438 U.S. 59, 85 (1978).

⁷³ *Id.*

⁷⁴ See Hugh W. Stephens, *The Texas City Disaster, 1947* (1997). The Texas City incident was the worst industrial accident in United States history, killing at least 581 people, injuring over 5,000 individuals and causing extraordinary amounts of property damage from ammonium nitrate blasts in the Port of Texas City. Following the incident, a class action was filed against the federal government under the Torts Claim Act, however the Courts refused to provide compensation for the victims because the Act may be invoked only on a "negligent or wrongful act or omission" of an employee, which created no absolute liability of the Government by virtue of its ownership of an "inherently dangerous commodity" or property, or of its engaging in an "extrahazardous" activity. After the court decision, Congress acted to provide compensation through Public Law 378, 69 Stat. 707 (1955). The last claim was processed in 1957, resulting in federal compensation of nearly \$17 million.

the public.⁷⁵

While upholding the constitutionality of the Act, the Supreme Court duly noted the legitimacy of and need for the cap on liability by stating “the limit on liability [is] ‘a classic example of an economic regulation—a legislative effort to structure and accommodate ‘the burdens and benefits of economic life.’”⁷⁶

The limit on liability remains the most controversial component of the Act, as critics argue that it constitutes a subsidy for the nuclear industry by not requiring unlimited liability. First of all, there is no record of the federal government ever paying a direct subsidy to any private licensees under Price-Anderson. The nuclear industry not only has paid the costs of the private, and secondary financial protection, insurance fees but it has also “paid millions of dollars in indemnity fees and has assumed more than \$9 billion in potential retrospective assessments to compensate injured accident victims – all of this at no cost to the government.”⁷⁷

In exchange for the limit on liability, the Price-Anderson Act provides a large, readily available source of compensation for any individuals affected from a nuclear incident that would otherwise not exist.⁷⁸ To the contrary, the Bhopal Disaster in India in 1984 demonstrates the problems with a system that fails to assure an available pool of funds in the event of an industrial accident, despite having no liability cap. The Bhopal Disaster is considered the world’s worst industrial catastrophe, as a leak of methyl isocyanate gas and other chemicals from a pesticide plant in Bhopal, Madhya Pradesh, India resulted in the exposure of hundreds of thousands of people to hazardous toxins.⁷⁹ The Indian government panel charged with tabulating the deaths and injuries determined that over 3,800 individuals died as a result of the leakage, 11,000 were disabled and an additional 150,000 to 600,000⁸⁰ were affected.⁸¹

Following years of litigation, the operating company, Union Carbide Corporation, settled with the Indian Government for \$470 million,

⁷⁵ *Duke Power* at 85-86 (quoting H.R. REP. NO. 89-883, at 6-7 (1965)).

⁷⁶ *Id.* at 83.

⁷⁷ See Quattrocchi, *supra* note 13, at 59.

⁷⁸ *Id.*

⁷⁹ Jackson B. Browning, *Union Carbide: Disaster at Bhopal*, in *CRISIS RESPONSE: INSIDE STORIES ON MANAGING IMAGE UNDER SIEGE* 365 (Jack A. Gottschalk ed., 1993).

⁸⁰ See AMNESTY INT’L, *CLOUDS OF INJUSTICE: BHOPAL DISASTER 20 YEARS ON* 61 (2004).

⁸¹ Browning, *supra* note 79, at 365.

approximately \$1,000 in compensation for each individual killed, disabled or injured from the disaster.⁸² The Price-Anderson Act represents not only a balancing of the risks of the nuclear industry but also of the protection of the American public. The truth of the matter is “that there is always a limit on liability – that limit equal to the assets of the company at fault.”⁸³ Those who drafted the Price-Anderson Act understood this and the legislative branch appropriately determined the private-public partnership, which established a liability threshold for the industry, was the most reliable system to ensure financial protection to the American people.

Throughout the five decades of the Price-Anderson Act, the public has never had to bear the economic brunt of any nuclear incident within the United States. Thus far, the insurance pools of the nuclear industry have paid more than \$200 million in claims and litigation costs since Congress passed the Act.⁸⁴ Out of this assessment, \$71 million in costs were disbursed following the Three Mile Island Accident in 1979.⁸⁵ The cost of nuclear commercial power plant insurance is borne by the industry, which is unlike various other energy sources within the United States. For example, the hydropower electricity industry is not responsible for incidents such as dam failure or resultant flooding; instead the public is the one to bear the burden of such costs.⁸⁶ This example is illuminated by the 1977 failure of the Teton Dam in Idaho, which caused approximately \$500 million in property damage, however the individuals affected from the failure were only compensated \$200 million of low-cost government loans.⁸⁷

In contrast, under the Price-Anderson Act, the insurance pools have absorbed \$200 million of the costs and the nuclear industry has paid \$21 million in indemnity fees to the federal government.⁸⁸ The success of this program has led Congress to extend the model to protect the public from other hazards or harm, such as medical malpractice, faulty vaccinations, toxic waste and terrorist attacks.⁸⁹ Congress should again adopt such a

⁸² See Quattrocchi, *supra* note 13, at 59.

⁸³ *Id.*

⁸⁴ Nuclear Energy Inst., *Price-Anderson Act Provides Effective Liability Insurance at No Cost to the Public*, (Jun. 2010), <http://www.nei.org/resourcesandstats/documentlibrary/safetyandsecurity/factsheet/priceandersonact/>.

⁸⁵ *Id.*

⁸⁶ GWYNETH CRAVENS, *POWER TO SAVE THE WORLD: THE TRUTH ABOUT NUCLEAR ENERGY 214* (Alfred A. Knopf 2007).

⁸⁷ *Id.*

⁸⁸ See Nuclear Energy Inst., *supra* note 84.

⁸⁹ *Id.*

model to extend towards clean energy technologies and production in order to properly encourage investment, development and innovation while also maintaining a high level of protection to the public.

B. NO-FAULT LIABILITY

In 1966, the legislative branch addressed concerns that many state tort laws required findings of fault or negligence in order to establish liability. This created a major obstacle for individual's seeking relief from the nuclear industry, as the technicalities and even knowledge of radioactive leakage, the nuclear industry and the proximate cause of an injury proved evasive. To appropriately resolve this issue, the Act implemented a waiver of defenses under state tort law in the event of a nuclear incident that shifted the standard essentially to one of strict liability.

Under this regime, claimants are legally required to only demonstrate that the injury or property damage sustained was caused by the release of nuclear material from the insured facility, however fault on a particular defendant does not have to be established.⁹⁰ The result of this provision is to effectively ensure a strict liability standard that provides the public with necessary protections from the judicial system. Such protections are essential in areas in which legal causation is difficult to prove.

C. FEDERAL JURISDICTION

State tort laws have historically governed nuclear liability determinations,⁹¹ however amendments to the Price-Anderson Act following the events at Three Miles Island revised the system in order to provide a federal overlay. Currently, the Act contains a pre-emption provision,⁹² which gives federal district courts jurisdiction over tort actions arising out of nuclear accidents and "expressly provides for removal of such actions brought in state court even when they assert only state-law claims."⁹³ The removal of such claims eliminates confusion and

⁹⁰ See Quattrocchi, *supra* note 13, at 57.

⁹¹ John L. Quattrocchi, *Nuclear Liability Insurance in the United States: An Insurer's Perspective*, in REFORM OF CIVIL NUCLEAR LIABILITY: INTERNATIONAL BUDAPEST SYMPOSIUM 1999 (OECD 2000).

⁹² 42 U.S.C. § 2014(hh) (2006).

⁹³ See *El Paso Natural Gas Co. v. Neztosie*, 526 U.S. 473, 484-85 (1999).

uncertainties surrounding the applicability of the Price-Anderson Act and establishes a level of assurance in how the judicial system will approach such claims.

Furthermore, as discussed in *El Paso Natural Gas Co. v. Neztosie*, the Price-Anderson Act “provides clear indications of the congressional aims of speed and efficiency.”⁹⁴ The chief judge of a district court is given the authority to appoint a special caseload management panel to oversee all filings and court hearings associated with a nuclear incident case.⁹⁵ These panels are designed to consolidate cases, set priorities, expedite cases or allow more equitable considerations of claims, and implement any measures as “as will encourage the equitable, prompt, and efficient resolution of cases arising out of the nuclear incident.”⁹⁶ Each of these provisions is in place to reduce the legal costs as well as promote efficiency and efficacy of the compensation process.

D. LIABILITY IS CHANNELED TO THE PARTICULAR LICENSEE RESPONSIBLE

The Price-Anderson Act channels financial responsibility and liability insurance obligations to the particular nuclear power plant responsible for the incident.⁹⁷ This mechanism helps assure that claimants will be provided financial compensation in the event of sustaining injury or property damage.⁹⁸ Under the Act, contractors, subcontractors, and suppliers to DOE contractors and NRC licensees, as well as the DOE contractors themselves, are fully indemnified for all liability.⁹⁹ These operators, however, are all connected with, or “channeled” to, a particular nuclear power plant. Accordingly, each power plant is responsible for indemnifying any accidents or incidents arising from its contractors, subcontractors, or suppliers activities. This is crucial in order to ensure full protection as well as development of the nuclear industry from all sectors. Without such assurance – both the economic assurance of indemnification for the public and legal insulation from individual liability for participating entities in the nuclear industry – the development of nuclear energy would certainly have faltered.

⁹⁴ *Id.* at 486.

⁹⁵ 42 U.S.C. § 2210(n)(3)(A) (2006).

⁹⁶ 42 U.S.C. § 2210(n)(3)(C)(vi) (2006).

⁹⁷ 42 U.S.C. §§ 2014(t), 2210(c) (2006).

⁹⁸ *See* Quattrocchi, *supra* note 13, at 56.

⁹⁹ Berkovitz, *supra* note 26, at 8.

E. GUARANTEED POOL OF FUNDS; CONTINUOUSLY WRITTEN

While criticism surrounds the liability cap of the Price-Anderson Act, the alternative of the Act is that the nuclear power plants would need to secure their own source of coverage. Not only, as mentioned already, would this detract most, if not all insurance companies, but in fact it would place the public in an extraordinarily unsettling situation. Establishing liability without the Price-Anderson Act would, in theory, place no legal limit on liability, however each claim would depend on state tort law and procedures, which may or may not provide for no-fault liability.¹⁰⁰ Even in the event that defenses are waived, a defendant with theoretically no liability limit might not be able to pay a judgment if obtained.¹⁰¹ Thus, the Price-Anderson Act establishes “assurance of prompt and equitable compensation under a pre-structured and nationally applicable protective system [which gives] way to uncertainties, variations and potentially lengthy delays in recovery.”¹⁰²

Under the Price-Anderson Act, compensation is evenly distributed over the entirety of those affected, however in an alternative system, such as a claim-based system, when the defendant’s assets are exhausted by earlier judgments, future claimants will be left without any compensatory relief or redress through judicial system.¹⁰³ Such a system would create an onslaught of lawsuits in order for claimants to be the first to express their grievances, rather than appropriately assuring the public a system that will orderly and equitably compensate those affected by any nuclear incident.¹⁰⁴

This sentiment was expressed in *Duke Power*, as the Supreme Court noted that

. . . the congressional *assurance* of a \$560 million^[105] fund for recovery, accompanied by an express statutory commitment, to “take whatever action is deemed

¹⁰⁰ See *H.R. 8631: To Amend and Extend the Price-Anderson Act Before Joint Comm. on Atomic Energy*, 94th Cong. 69 (1975) (statement of William A. Anders, Chairman, U.S. Nuclear Regulatory Commission).

¹⁰¹ See Quattrocchi, *supra* note 13, at 59 (“The simple fact is that there is always a limit on liability—that limit equal to the assets of the company at fault.”).

¹⁰² *Duke Power Co. v. Carolina Envtl. Study Grp*, 438 U.S. 59, 89 (1978).

¹⁰³ *Id.* at 90.

¹⁰⁴ *H.R. 8631: To Amend and Extend the Price-Anderson Act*, *supra* note 101.

¹⁰⁵ This was the figure at the time of *Duke Power*, however, the current assurance from both the primary insurance and secondary insurance pool is \$12.6 billion.

necessary and appropriate to protect the public from the consequences of” a nuclear accident, [is] a fair and reasonable substitute for the uncertain recovery of damages of this magnitude from a utility or component manufacturer, whose resources might well be exhausted at an early stage.¹⁰⁶

Furthermore, the American Nuclear Insurer policies are written on a continual basis and contain no expiration date.¹⁰⁷ Claims based during the time of the policy are accounted for and the licensee is still held responsible if the claim is found valid, even if they are no longer in operation.

F. LITIGATION AND INVESTIGATION COSTS INCLUDED

Under the Price-Anderson Act, the expenses of investigating and defending claims or suits against the nuclear industry are included in the limit of liability.¹⁰⁸ The legal costs for defending many of these actions can be quite expensive. By including the legal and investigation costs in the Act, Congress established definite confines for liability costs that insurance companies providing the financial protection plans to the nuclear industry could rely upon. In essence, the inclusion of these costs enables insurers to offer their maximum capacity commitments without fear of exceeding such commitments.¹⁰⁹ This provision is crucial in enabling insurers to maintain and, most likely, increase the assets they place at risk.¹¹⁰

G. THE PRICE-ANDERSON ACT PROVIDES STABILITY IN THE MARKET

Finally, the requirement of the Price-Anderson Act that each nuclear commercial power plant must obtain a specified amount of private insurance as well as participate in the secondary financial protection program provides stability in the market that might otherwise not be there. Not only was the private insurance industry precarious in providing financial protection for the nuclear industry at the dawn of its development, but the liability insurance market is also a volatile entity by nature.

¹⁰⁶ *Duke Power Co.*, 438 U.S. at 90-91 (citation omitted).

¹⁰⁷ Richard Jones, *Nuclear Insurance: Where Does it Fit in the Green Generation?*, 16 J. REINSURANCE 71, 77 (Spring 2009).

¹⁰⁸ See 42 U.S.C. § 2210(e) (2006).

¹⁰⁹ See Quattrocchi, *supra* note 13, at 56.

¹¹⁰ *Id.*

This very instability was demonstrated by the liability insurance crisis from late 1984 through 1986, when major economic disruptions in the commercial liability insurance market created concerns over the availability and affordability of a number of commercial insurance policies – notably for chemical and pharmaceutical companies, the medical system and municipalities.¹¹¹ However, when the liability insurance crisis hit the nation in the mid-1980s, the nuclear liability insurers continued to provide a stable market for their limited customer base as a result of the system provided by the Price-Anderson Act.¹¹² The nuclear industry was shielded from any increase in liability premiums, cancellation of policy coverage or diminishment in scope of coverage¹¹³ – not only was this critical for the nuclear industry but it also protected the public from any exposure to an uninsured, or volatile, nuclear industry.

VII. CARBON SEQUESTRATION: AN OVERVIEW

With the growing demand and development of carbon sequestration technology, insurers are beginning to determine the best approach in providing liability frameworks for private, state and federal projects using this new technology. Briefly, carbon sequestration is the process that involves capturing carbon dioxide at the point of combustion – most notably from coal power plants – and injecting it into geological formations beneath the surface of the earth.¹¹⁴ Essentially, the technology is the reverse of pumping oil or natural gas from a confined geological aquifer.¹¹⁵ The life-cycle of a carbon sequestration project can last over a couple of centuries as the process involves several phases from site selection, characterization, and regulatory review;¹¹⁶ to CO₂ injection and well closure;¹¹⁷ to post-closure monitoring;¹¹⁸ and finally to long-term

¹¹¹ G.C. Lai et al., “*On Liability Insurance Crisis*”, Risk Theory Seminar Conference, Univ. of Ala. 1 (Apr. 1997).

¹¹² See Quattrocchi, *supra* note 13, at 58.

¹¹³ See Lai, *supra* note 111.

¹¹⁴ INT’L ENERGY AGENCY, TECHNOLOGY ROADMAP: CARBON CAPTURE AND STORAGE 8-9 (2009).

¹¹⁵ Alexandra B. Klass & Elizabeth J. Wilson, *Climate Change and Carbon Sequestration: Assessing a Liability Regime for Long-Term Storage of Carbon Dioxide*, 58 EMORY L.J. 103, 115 (2008).

¹¹⁶ *Id.* This phase is expected to last anywhere from one to ten years.

¹¹⁷ *Id.* (twenty to thirty year life-span).

¹¹⁸ *Id.* (estimating that this phase of the life-cycle lasts for a period of fifteen to thirty years).

stewardship.¹¹⁹

To date, carbon dioxide is the most abundant anthropogenic greenhouse gas in the atmosphere largely due to human activities.¹²⁰ The release of these gases into the atmosphere has contributed to global warming and increasing concerns of altering climatic, biological, and natural environments. The demand, and necessity, for solutions to mitigate greenhouse gas emissions are the propellant behind carbon sequestration projects.

While the technology is only increasing, carbon capture and storage projects are technically ready – but the associated costs, including insurance, need to be lowered and investment needs to increase in order for large-scale implementation of this technology.¹²¹ In essence, the industry today parallels the circumstances of the nuclear industry in the 1950s, when the Atomic Energy Act was originally enacted. The manner in which the insurance industry approaches the long-term liability with carbon sequestration could significantly affect the development and investment in the technology.

¹¹⁹ *Id.* (estimating that CO₂ remains sequestered underground for hundreds of years).

¹²⁰ Carbon dioxide is the most important anthropogenic greenhouse gas (*see* Figure SPM.2). The global atmospheric concentration of carbon dioxide has increased from a pre-industrial value of about 280 ppm to 379 ppm in 2005. The atmospheric concentration of carbon dioxide in 2005 exceeds by far the natural range over the last 650,000 years (180 to 300 ppm) as determined from ice cores. The annual carbon dioxide concentration growth rate was larger during the last 10 years (1995–2005 average: 1.9 ppm per year), than it has been since the beginning of continuous direct atmospheric measurements (1960–2005 average: 1.4 ppm per year) although there is year-to-year variability in growth rates. The primary source of the increased atmospheric concentration of carbon dioxide since the pre-industrial period results from fossil fuel use, with land-use change providing another significant but smaller contribution. Eleven of the last twelve years (1995–2006) rank among the 12 warmest years in the instrumental record of global surface temperature (since 1850). *See* INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2007, *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (S. Solomon et al. eds., Cambridge University Press 2007).

¹²¹ *Id.*

VIII. CARBON SEQUESTRATION RISK ASSESSMENT & COMPARISON TO THE NUCLEAR INDUSTRY

In regards to insurance necessities, policy writers must first understand the risks in which the technology presents. The inquiry first begins with what is to be protected and what threshold of risk mitigation required or desired by the client.¹²² In general, insurance policies are structured, priced and conditioned based on the frequency and the severity of potential loss.¹²³ In developing such policies, insurers look towards past events and historical trends within the insured's field and practice area.

The difficult part of carbon sequestration is that the infancy of the technology does not lend itself to any insight in its past history. As a result, the insurance industry has reluctantly provided coverage for the risks and trends of the technology. However, this should not detract insurers, as the underlying liabilities and risks of carbon sequestration technology are analogous to the same uncertainties that faced the nuclear industry in the 1950s.

First, the main concern with carbon sequestration is the release of carbon dioxide from the project site. In its natural state carbon dioxide is non-toxic, however concentrations of 5-10% by volume is harmful to the life and health of plants, humans and animals.¹²⁴ Thus, there is a direct and measurable potential for damage resulting from the release of large quantities of carbon dioxide in the event of a site failure or storage site leakage due to a number of reasons such as equipment or construction failure, unexpected tectonic movements or unforeseen large-scale migration.¹²⁵ The inherent risks associated with the storage of carbon dioxide are similar, although much smaller, to radioactive exposure resulting from a nuclear power plant failure.

Secondly, there is general consensus that potential leakage of hazardous pollutants from storage reservoirs is very low – the Intergovernmental Panel on Climate Change estimates that for well-selected sites, there is a 90-99% probability that over 99% of liquefied CO₂

¹²² PA. DEP'T OF CONSERVATION AND NATURAL RES., ASSESSMENT OF RISK, LEGAL ISSUES, AND INSURANCE FOR GEOLOGIC CARBON SEQUESTRATION IN PENNSYLVANIA 5-1 (2009).

¹²³ *Id.*

¹²⁴ Christina Ulardic, *Environmental Impairment Liability Insurance for Geological Carbon Sequestration Projects*, in INTERNATIONAL RISK GOVERNANCE COUNCIL WORKSHOP ON REGULATION FOR CCS 1, 3 (Mar. 2007), available at http://www.irgc.org/IMG/pdf/IRGC_CCS_SwissRe07.pdf.

¹²⁵ *Id.*

injected into underground wells will remain underground for over 100 years.¹²⁶ Again, this is analogous to the nuclear industry's low-frequency yet high-risk nature.

Also, while immediate injuries and harm can arise from industrial incidents with carbon sequestration at the time of the failure, the associated liability is often a result of storage leakage. In turn, this leads to injuries and harm that are latent in nature, as leakages tend to arise over long-term, chronic exposure to low-levels of hazardous materials. Likewise, a radiological incident can result in acute, short-term exposure as well as chronic, long-term exposure to surrounding communities and environments.¹²⁷ In both cases, victims are hindered with the difficulty of proving causation arising from both nuclear incidents and carbon sequestration failures. Finally, carbon sequestration is beginning to solidify itself as a crucial part in our national energy policy similar, albeit on a smaller scale, to the nuclear industry's rapid ascent to the forefront of the energy sector in the middle of the twentieth century.¹²⁸ This commitment from the federal government intensifies the need to create an indemnity scheme that can serve the dual purpose of promoting the technology while also protecting the public.

Admittedly the release of radioactive materials is far more damaging to the public than concentrated volumes of carbon dioxide,

¹²⁶ INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE, CARBON DIOXIDE CAPTURE AND STORAGE 14 (Bert Metz et al. eds., 2005).

¹²⁷ NAT'L CTR. FOR ENVTL. HEALTH CTRS. FOR DISEASE CONTROL & PREVENTION, THE PUBLIC HEALTH CONSEQUENCES OF DISASTERS 405 (Eric K. Noji ed., Oxford Univ. Press 1997) (Acute effects from radiological exposure vary in dosage from individual showing no outward symptoms, but instead having increased chromosomal aberrations in blood lymphocytes and lower blood count to high doses, which may affect the central nervous system causing seizures, gait disturbances and coma, almost always resulting in death. However, this high-dosage acute exposure is extremely rare in nuclear incidents; instead it is seen in intentional nuclear warfare.).

¹²⁸ See American Recovery and Reinvestment Act of 2009, 26 U.S.C. § 46 (2009) (The Advanced Energy Manufacturing Tax Credit, also known as Internal Revenue Code Section 48C, provides a 30% tax credit for future expenditures to support new, expanded, or re-equipped domestic manufacturing facilities for advanced energy projects. The tax credit was promulgated pursuant to the American Reinvestment and Recovery Act (ARRA) section 1302, which authorized the Department of Treasury to extend \$2.3 billion for qualified investments in domestic manufacturing facilities that can be completed within a four year period. Credits are available for a two year period, or until the maximum dollar amount of credits has been reached.).

however the risk that is being covered by both industries mirror each other in associated problems – leakage of material causing exposure to individuals; groundwater contamination; wind-blown migration of the elements; and property damage due to an industrial meltdown or explosion. The Price-Anderson Act has already accounted for the similarities in the problems associated with each industry through, among other provisions, its liability channeling, no-fault causation standard, and removal to federal jurisdictions.

In addition, the private-public partnership of the Price-Anderson Act provides suitable incentives such as a liability cap, industry-wide pooling, and policies written on a continual basis that would attract greater financial investment, insurance coverage and ultimately employment of carbon sequestration technology.

A. NECESSITY OF A LIABILITY CAP

First of all, the liability cap is essential in attracting private insurance interest in new, clean energy sources that have yet to reach maturity within the market. This was an essential component for the nuclear industry, as the potential liability costs associated with a nuclear incident detracted private insurance companies from providing coverage for nuclear power plants. Likewise, the long-term risks associated with carbon sequestration concern insurance underwriters today. Thus, there needs to be incentives for private investment to generate the necessary development of the field – the liability cap is the first step in doing so. As was the case with the Price-Anderson Act, the federal government should implement a preliminary system that caps site-specific liability at a certain threshold while also ensuring federal indemnification in the event of an incident exceeding the primary, site-specific insurance. Then, once the industry begins to mature, the federal government can slowly turn the second-tier of the indemnification system over to private insurers that begin to adapt to and draw interest in the technology.

The tiered-system that blends site-specific insurance and an industry-wide pool of funds “both provide[s] site-tailored risk management and ensures[s] that adequate funds are available to cover damage in the post-closure period.”¹²⁹ In addition, an industry-wide pool would allow for risk sharing on a national scale and financial protection for a variety of different projects. If indemnification was instead based on a state-level then proper financial protection might quickly be exhausted without the same

¹²⁹ See Klass & Wilson, *supra* note 115, at 167.

amount of diversity, contribution and risk sharing that occurs in a tiered, mutual insurance pool.¹³⁰

Federal involvement is even more crucial with carbon sequestration because the technology does not tend to yield intensive amounts of revenues. In fact, the main concern of the technology is in diminishing the amount of carbon emissions in the atmosphere, not in turning a profit. Therefore, if there is to be individual and private investment in such projects, monetary or regulatory incentives must be in place for further development.¹³¹

B. REMOVAL OF CLAIMS TO FEDERAL COURT

The American Recovery and Reinvestment Act of 2009 was signed into law by President Obama on February 17, 2009 and provided the Office of Fossil Energy with \$3.4 billion in the attempt to fund initiatives focused on research, development and deployment of technologies to use coal more cleanly and efficiently.¹³² In order to accomplish these goals, federal agencies have invested significantly in carbon capture and storage projects across the nation. When dealing with issues such as our national energy policy, federal courts should be the forum for settling disputes, rather than in a plethora of state courts, especially if the technologies are targeted to decrease *national* carbon emission standards. Such an approach maintains cohesiveness as well as ensures an unbiased, equitable forum. The Price-Anderson Act contains a preemption provision¹³³ that not only gives federal courts jurisdiction over tort action arising out of nuclear accidents, but it also provides for removal of such actions brought in state court even when they assert only state-law claims.¹³⁴ This provision is essential for carbon sequestration claims.

First, the provision will limit the consequences of arriving at different conclusions on the applicable law as a result of a inherent differences throughout jurisdictions and variances in state tort law. Second, the tension between state law and federal preemption is a constant theme for CCS, especially due to the potential damages occurring in

¹³⁰ *Id.* at 167-68.

¹³¹ See ULARDIC, *supra* note 125, at 2.

¹³² The Nat'l Energy Tech. Lab., *Carbon Sequestration: American Recovery and Reinvestment Act Funding (AARA)*, http://www.netl.doe.gov/technologies/carbon_seq/arra/index.html (last visited Sept. 6, 2011).

¹³³ See 42 U.S.C. § 2210(n)(2) (2006).

¹³⁴ See *El Paso Natural Gas Co. v. Neztosie*, 526 U.S. 473, 484 (1999).

domains with strong state laws governing groundwater protection, mineral rights, or surface rights.¹³⁵ Carbon sequestration claims, however, should be heard under a federal overlay due to the fact that a lot of the water and mineral resources that will be in question involve several states as well as multiple state interests such as agriculture, urban development, tax revenues and wildlife preservation.¹³⁶ Thus, an indemnity system must include a provision that removes any claims from state court in order to preserve all interests in the matter, provide for efficiency and efficacy, and establish cohesive law regarding carbon sequestration facilities and events.¹³⁷

Federal jurisdiction eliminates many of the uncertainties individual victims might encounter in state courts, such as more stringent causation standards or heightened burdens of proof. Additionally, the carbon sequestration liability scheme should adopt similar provisions to the Price-Anderson Act that give the chief judge of the federal district court the authority to consolidate cases, set priorities, expedite cases or allow more equitable considerations of claims, and implement any measures as that will encourage the equitable, prompt, and efficient resolution of cases arising out of the nuclear incident.¹³⁸ Such provisions will expedite compensation for affected individuals, which is in essence, the entire objective of the system.

C. CAUSATION: NO-FAULT IS GOOD

In order to provide proper protection to the public in the case of a storage leakage, the insurance policy should be written in such a manner that establishes a strict liability standard. Due to the long duration of carbon storage facilities, it will be much more difficult to detect and assign responsibility for any harm that might occur.¹³⁹ Furthermore, the latent nature of injury or property damage associated with carbon storage leakage hinders individual claimants from providing any other standard of legal causation.

¹³⁵ See Klass & Wilson, *supra* note 115, at 168.

¹³⁶ *Id.*

¹³⁷ *Id.*

¹³⁸ 42 U.S.C. §§ 2210(n)(3)(C)(i), (ii), (v), (vi) (2006).

¹³⁹ SEAN MCCOY, CARNEGIE MELLON UNIV. DEP'T OF ENG'G & PUB. POLICY, POLICY BRIEF: COMPENSATION, LIABILITY AND LONG-TERM STEWARDSHIP FOR CCS 4 (July 13, 2009), http://www.ccsreg.org/pdf/LongTermLiability_07132009.pdf.

In addition, the judicial system has already established a long history of strict liability standard for torts similar to carbon sequestration leakages. This principal stems from *Rylands v. Fletcher*, the English decision which held that an individual “who for his own purposes brings on his lands and collects and keeps there anything likely to do mischief if it escapes, must keep it in at his peril, and, if he does not do so, is prima facie answerable for all the damage which is the natural consequences of its escape.”¹⁴⁰

This body of tort law has been extended in United States laws regarding cases of abnormally dangerous activities, and also, hazardous materials. Thus, in order to maintain consistency with legal precedent as well as ensure that affected individuals are duly compensated, a no-fault liability system is essential in carbon leakage claims.

D. ADDITIONAL ARGUMENTS FOR PRICE-ANDERSON ACT

The Price-Anderson system contains several other provisions that would be highly beneficial for the carbon sequestration system to incorporate. First, the Act channels liability to the particular facility that is responsible for any industrial incident. Carbon sequestration technology has the potential of including a large network of individuals due to its operating, post-injection and long-term stewardship phase. Thus, a large number of individuals are exposed throughout the phases of the technology, creating issues of widespread liability. Thus, an “omnibus” liability system, which allows for all suppliers, transporters, and participants of the carbon sequestration industry to be covered under the insurance system would not only be instrumental for the industry, but also for the public as an “omnibus” feature “permits a more unified and efficient approach to processing and settlement of claims.”¹⁴¹ Additionally, this establishes a centralized defendant in the event that affected individuals are unsure as to the negligent, or directly responsible party.

Second, the Price-Anderson Act provides a mutual insurance program that guarantees a certain amount of available funds in the event of an incident. This system provides for compensation to be evenly distributed

¹⁴⁰ *Fletcher v. Rylands*, [1866] 1 L.R. Exch. 265 at 279 (Eng.).

¹⁴¹ OMER F. BROWN, II, ENERGY CONTRACTORS PRICE-ANDERSON GRP., REPLY COMMENTS ON U.S. DEPARTMENT OF ENERGY NOTICE OF INQUIRY CONCERNING PREPARATION OF REPORT TO CONGRESS ON THE PRICE-ANDERSON ACT 4 (Dec. 31, 1997), http://energy.gov/sites/prod/files/gcprod/documents/REPLY_PAGroup.pdf.

to all affected individuals, while also assuring federal indemnification when necessary. This is more functional than having multiple state systems, which may not sufficiently spread risk and may not be adequately capitalized to cover actual damages incurred above insurance coverage limits.¹⁴² Also, the Price-Anderson Act policies are written on a continual basis, ensuring the public that regardless of the timing of an incident the policy will cover any injuries, harm, or property damage so long as the claims are filed within a timely manner as provided for by the statute of limitations.

Finally, the Price-Anderson system includes litigation and investigating costs. This is crucial in order for the insurance industry to be fully aware of the risks that they are taking. Companies are able to properly assess the amount of financial protection they want to provide, while also allowing them to secure reinsurance to spread the risk over an even greater base.

IX. GEOTHERMAL ENERGY

In the past couple of years, investor interest in geothermal technology has increased significantly. This rapid increase in investment has been accelerated by growing demand for energy sources, increases in the price and scarcity of oil, and the developing awareness of the risks presented by carbon emissions.

With Enhanced Geothermal Systems (EGS), a well known as a production-injection well is drilled into hot basement rock that has limited permeability and fluid content.¹⁴³ Hot, dry rock that is closer to the earth's surface is ideal for this technology. The production-injection well consists of two drill points, the first being the injection well that serves to pump water under high pressure into the earth's core. Pumping the water under high pressure is to ensure fracturing or increase fracturing within the geological environment, thus creating an artificial geothermal reservoir.¹⁴⁴ Water is then circulated through the reservoir and the hot water is extracted

¹⁴² Charles H. Haake & Karyn B. Marsh, *Climate Change: Carbon Sequestration*, WORLD CLIMATE CHANGE REPORT, May 8, 2009, at 1, 6.

¹⁴³ U.S. Dep't of Energy, *The Basics of Enhanced Geothermal Systems*, GEOTHERMAL TECHNOLOGIES PROGRAM, http://www1.eere.energy.gov/geothermal/pdfs/egs_basics.pdf (last visited Sept. 6, 2011).

¹⁴⁴ Kai Sametinger, *How to Invest in Geothermal*, RENEWABLE ENERGY FOCUS, Jan.-Feb. 2009, at 84, available at <http://www.renewableenergyfocus.com/view/886/investing-in-geothermal/>.

from the production well, which is drilled with the intent to intersect the stimulated fracture system created by the injection well in a manner in which the most amount of the artificial reservoir is in contact with the well. In turn, the water extracted, known as brine, is pumped through an electrical power plant and the brine heats a working fluid that produces vapor to drive a turbine-generator.¹⁴⁵ The original water is then recycled through a cooling facility and is re-injected into the reservoir, thus completing the cycle.¹⁴⁶

With the increase in investment in geothermal interest and the growing development of geothermal energy facilities, policy providers must be concerned and appropriately assess the inherent technical perils in the testing, construction and maintenance of these geothermal facilities. Most geothermal power projects take five to seven years to be operational, as each phase of the project has its own set of requirements and risks attached.¹⁴⁷ This concern is further exacerbated by installation, operation and development of projects in harsh unstable terrain, proximity to marine environments, and drilling necessities.

The current investment market for geothermal technologies is relatively weak. The length of the projects combined with the nascent history of the technology deters investors of these projects. The success of drilling – determined by the volume, temperature and pressure of the fluids discovered – is crucial to the financial stability of the project, as it consists of up to 30-40% of the entire project.¹⁴⁸ Therefore, in some cases government support and subsidies are necessary to help get the project off the ground.¹⁴⁹ Investors that then have to worry about insurance costs associated not only with the building but also with the maintenance and operation of the project are further deterred, as additional costs must be accounted for. Therefore, the Price-Anderson Act, which would provide incentive for financial investment by implementing liability caps in the event of an incident, failure in project development, and other instances, is crucial to incentivize financial investment.

¹⁴⁵ U.S. Dep't of Energy, *supra* note 143.

¹⁴⁶ *Id.*

¹⁴⁷ Sametinger, *supra* note 144, at 85.

¹⁴⁸ *Id.*

¹⁴⁹ *Id.*

X. CONCLUSION

The rapidly rising demand for electricity, increasing costs of oil and gas, and concerns about energy security demonstrate the need not only for the “nuclear renaissance,” but also for investment in alternative, clean energy technologies and sources.¹⁵⁰ The past few years have illuminated a heightened national interest and political commitment, from both sides of the aisle, regarding these objectives.

If the United States expects to continue to be atop the global pyramid, then it must reassess its current energy policy – especially its commitments to nuclear energy as well as alternative, clean energy technologies. Currently, the nuclear industry accounts for 19.4% of electrical production within the United States, while at the same time accounting for 73.6% of the emission-free electricity production.¹⁵¹ In 2006 alone, the nuclear industry saved the United States and the world 681.2 million metric tons of CO₂ emissions while providing the lowest-cost producer of base-load electricity at 1.72 cents per kilowatt-hour.¹⁵² However, the United States’ nuclear production is ninth in the world in percent of its total domestic electricity generation.¹⁵³

Clean energy technologies, such as nuclear energy, are at the forefront of the national discourse on energy policy, while others, such as solar and wind power, geothermal developments, and carbon sequestration, are being pushed into the discussion. It is in these developments that the proper financial instruments and insurance policies must adequately support and protect this increasing development.

Accordingly, analysis of the Price-Anderson Act provides insight into an underlying insurance indemnification system that would provide a well-suited framework in addressing the growth of clean energy technologies, such as geothermal energy and carbon sequestration technology. These and similar technologies are currently struggling for the necessary financial backing and protection in order to become a key contributor to our national energy policy. As such, the liability cap

¹⁵⁰ HOUSE SELECT COMM. FOR ENERGY INDEPENDENCE & GLOBAL WARMING, 110TH CONG., THE REALITIES OF NUCLEAR EXPANSION 2 (Sharon Squassoni 2008).

¹⁵¹ Roger Bloom, *The Advantages of Nuclear Power in a Carbon Constrained World*, WESRCH.COM 1, 27-28 (Apr. 27, 2009), <http://energy.wesrch.com/paper-details/pdf-TR17TE000CHAC-the-advantages-of-nuclear>.

¹⁵² *Id.* at 29-30.

¹⁵³ *About ANI*, AM. NUCLEAR INSURERS, <http://www.amnucins.com/About%20ANI.html> (last accessed Sept. 7, 2011).

instituted under the Price-Anderson Act is essential in encouraging growth within the clean energy sector.

Likewise, a no-fault system in which liability is channeled to the particular operator or owner of the facility is crucial in upholding accountability and protection to the American public. Finally, once a national market is established, creating a tiered system that requires site specific private insurance, a secondary, industry-wide pool, and finally federal indemnification in the event of exhaustion of the initial two layers will create a reliable network of compensation in the event of an industrial incident while also balancing the competitive needs of the clean energy sector. In addition to these components, future clean energy indemnification systems should include written insurance policies; inclusion of litigation and investigation costs in the liability limitations; and removal and consolidation of all claims to one federal district court.

Clean energy technologies are the future in American power production. Such technologies offer a number of similarities – most notably the inherent nature of high-impact, low frequency risk in their development, maintenance and production processes. Accordingly, these energy sources invite a more in-depth analysis into the best practices in promoting and developing the proper technological advancements, financial investments, and insurance policies to protect and promote the development of clean energy while safeguarding the American public. The Price-Anderson Act should be at the forefront of this analysis.

